2007

Color vision screening using eye movements

Natthapongs Voraphani
Iowa State University

Color vision screening using eye movements

by

Natthapongs Voraphani

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Industrial and Agricultural Technology

Program of Study Committee:
Brian L. Steward, Major Professor
Derrick J. Parkhurst
Steven A. Freeman
Lie Tang
Richard B. Evans

Iowa State University

Ames, Iowa

2007

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1.

CHAPTER 2.

CHAPTER 3.

CHAPTER 4.

LIST OF TABLES

ACKNOWLEDGEMENTS

I sincerely thank my major professor, Dr. Brian Steward for providing me with the opportunity to pursue my PhD here. I value your guidance and thank you for allowing me to think freely and for helping me find ways to go through with my research. I feel grateful for your kindness. I would like to thank my committee members, Dr. Steven Freeman, Dr. Lie Tang, and Dr. Richard Evans for all the advices and suggestions. I have enjoyed interacting with each of you. I have learned much about occupational safety as well as educational research with Dr. Freeman. I admire Dr. Evans for his innate curiosity in statistical science and how he can bridge the research knowledge to other fields. I thank Dr. Tang for keeping me on track of my study and for all the valuable suggestions he provided me.

I would not have accomplished this work without the help from Dr. Derrick Parkhurst. I cannot adequately thank you for all the guidance you have provided to me both professionally and personally. Throughout the years that I have been a part of your human computer vision lab, I was fortunate to have seen and learned how to act as a creative thinker and critical learner at the same time from you. I will never forget how tirelessly you have worked to help your students improve themselves. And without your guidance, all of this research work would not have happened.

My deepest appreciation goes to my beloved parents who I am so proud of, Pairuch and Aurawan Voraphani, you both always believed in me and supported my dreams. My special thanks to my mother, for always looking over me and for making sure I had everything and was happy at all times. I thank my little sister and two brothers, Nipasiri, Vuttichai, and Paisit for the moral support and for everything you guys have helped me across the seas to cheer me up and keep me going. I thank my brother-in-law, Kittiyut and Geoff Wild also for providing me with all those get away days to San Jose. Those days will always be treasured in my mind. I thank the Kitikoon family, Dr. Viroj (Pa Roj) and Dr. Tuanta (Aa Tuan) for accepting me and all the kind support from both of you through these years.

Finally, my last thanks must go to Pravina, my wife. Her love and support were absolutely indispensable. Her kindness, patience, and caring for me were truly above and

beyond anything I expected. I feel very fortunate I have a life's partner with an intuitive understanding of how hard it would be when both of us wanted to pursue a Ph.D. degree abroad. I felt lucky we, with our Cartoon, were strong enough that we could go through when things just didn't go our way. For this success, I am immensely grateful and deeply indebted to her. This dissertation is dedicated to her.

ABSTRACT

Most of color vision tests require that the participant is aware of and actively takes part in the test. While this might seem like a minor requirement, it presents the possibility that, when motivated, participants can engage in subterfuge in order to pass the test. This study introduces the use of eye movements as a covert test to detect the presence of a color deficiency without requiring explicit input. An experiment was conducted in which the results suggest that the observed differences in eye movements alone can be used as an efficient metric for detecting colorblindness using the Ishihara images. To reduce the possibility that participants may realize they are being tested, a method to select new complex scenes for screening color vision status was developed. Another eye movement experiment was performed using selected scenes to validate the usefulness of the proposed method. The results show that this approach can be used to screen colorblind individuals without their knowledge or active participation, making the test less sensitive to subterfuge.

CHAPTER 1. GENERAL INTRODUCTION

1.1 Statement of Problem

The ability to perceive colors is one of the most important functions of human vision. Color vision not only enhances human perceptual capabilities but also serves as a visual cue for daily life activities. In fact, many professional activities depend on color discrimination ability for both job performance and safety. However, colorblind individuals lack the ability to discriminate certain colors. Hence, colorblind individuals will tend to confuse the meaning of information that use those colors. Worse, some colorblind individuals do not realize that they have a color deficiency until required to perform tasks or undertake occupations that require color judgments [1]. Such a risk is illustrated by the following example.

In July 2002, a Federal Express Boeing 727 plane struck trees and crashed short of an airport runway at Tallahassee, Florida. The National Transportation Safety Board investigated the accident and found that the pilot's inability to identify the correct colors of the path indicator was one of the major causes of the accident. Surprisingly, it was reported that the pilot had passed the color vision battery test required by the Federal Aviation Administration (FAA) regulation. The board made a suggestion to the FAA that existing color vision tests may not provide adequate detection of color vision deficiencies in some situations where the speed and accuracy of color recognition is critical to work function [2]. This example demonstrates two points. First, color vision deficiencies can adversely affect job performance. Second, conventional color vision testing protocols may be of limited use for certain critical tasks in particular professions, due either to lack of sensitivity or the potential that individuals may engage in subterfuge to pass the test.

The evidence from several studies suggests that career choice and job performance can be adversely affected by color vision deficiencies [1, 3-7]. Individuals with severe color vision defects are slower and make more errors recognizing signals, including maritime, aviation, rail, and road signals as compared to normal observers [8-12]. They also have problems discriminating natural surface colors (e.g., fluids, tissues, body skin, foliage) [4, 13] as well as man-made surface colors (e.g., wires, color display, chemical substances, art prints) [14-16]. Individuals with color vision deficiencies also report difficulties with everyday color related activities [17]. These include, for example, selecting colors of clothing or paint, identifying colors of flowers or ripeness of fruit, and recognizing colors on a television or computer screen.

Occupational organizations have issued color vision requirements for some professional groups due mainly to the high importance of color vision in their work activities, and the fact that the social, environmental and economic cost of job related mistakes and accidents can be very high [1, 6]. Some voluntary guidelines have clearly emphasized normal color vision as a necessary condition to perform essential job functions safely. For example, the FAA has required airmen to provide a medical certification for airline employment, including color vision test results [2]. A consensus standard of the American Society of Mechanical Engineers requires that operators of crane machines be able to distinguish between colors. Although the Occupational Safety & Health Administration (OSHA) does not require normal color vision to perform any particular job function, many occupational associations have standardized their own specifications in relation to the relevant visual tasks [18]. It is recommended by the Committee on Vision, National Research Council, that individuals performing occupations that require color discrimination be screened for color

deficiencies in order to assure occupational safety [19]. These include, but are not limited to, careers in the fields of marine navigation, civil aviation, railway operation, hospital laboratory, construction, armed services, transportation, fire services, textile manufacture, photography and fine art industries [1, 20-22].

Several diagnostic tests have been developed to detect deficiencies in color vision. The most widely used test is the Ishihara color vision test as it has a high degree of sensitivity and specificity [23]. The Ishihara test consists of a series of images that depict numerals using arrays of small colored dots. The dots in the array that make up the numeral itself are all of similar color (e.g., red). The remaining dots that make up the background and are all of a similar color (e.g., green) that is different from the color of the numeral. Individuals with normal color vision perceive these colors differently and can thus detect the numerals that are portrayed by the pattern of dots. On the other hand, individuals with color deficiencies can have trouble differentiating these colors and often fail to detect the numerals. The Ishihara test exploits this fact and asks participants to view a series of such images and to guess the identity of the numerals that are present in the images. Although this test can detect the presence of a color deficiency, it does not indicate the type of deficiency [24]. Moreover, since the test mainly uses numerals, it is not a viable test for young children and people who cannot understand or read numerals [25].

Another color vision test is the Farnsworth-Munsell 100-hue test (FM-100). This test utilizes a set of 85 colored caps. The caps are arranged in four cases. For each case, participants are asked to arrange these caps next to reference caps so that the colors are sequentially ordered in a circle. The total number of ordering errors can be used to detect the presence of a color deficiency. However, the FM-100 test is significantly more time

consuming to conduct than the Ishihara test. This test can take as long as 20 minutes while the Ishihara test can be completed in 3 minutes [26, 27].

A modification of the FM-100 test which involves the arrangement of 15 colored caps is called the Farnsworth Panel D-15 Test. This test classifies participants into two groups: those with normal color vision and those with strong/medium color defects. Although the D-15 is a quick screening test that allows 2 minutes to arrange the caps, many practitioners reported that the D-15 is extremely difficult to administer to small children and elderly people [19, 28].

A clinical instrument that can be used to diagnose all types of color vision deficiencies is the anomaloscope. The anomaloscope diagnoses red-green defects by assessing the observer's ability to match a pair of red and green wavelengths. The participant is asked to look down a viewfinder and adjust the proportions of red and green in a mixture field until the mixture appears identical to a yellow color in a test field. However, this test is quite expensive, complicated to implement, time consuming, and requires specialized training [27, 29].

All of these tests can reliably detect the presence of a color deficiency, but they require that the participant is aware of and actively takes part in the test. While this might seem like a minor requirement, it presents the possibility that, when motivated, participants can engage in deception in order to pass the test [12]. For example, the introduction of color vision screening into the workplace for occupational safety reasons may lead colorblind individuals to memorize the correct responses to the test in order to keep their jobs [4]. It is already established that performance on two of the most prevalent color vision tests, the

Ishihara and the Farnsworth-Munsell 100-hue test, can be improved with familiarity or training [4, 12].

While modification to the most prevalent color vision tests may help limit subterfuge, most of these increase the amount of participation required in form of verbal communication or manual input. Therefore, these tests can only be applied with difficulty to some groups such as elderly adults or individuals with disabilities. Some participants realize that the testing procedure is difficult and become embarrassed or impatient with the test [19]. Some variants of the arrangement test are difficult to administer unless the participant fully understands the task [24]. Color vision tests that require complicated tasks or instructions can easily confuse participants [19]. An alternative method that covertly screens for color deficiencies by not depending on the explicit response of the participants would limit subterfuge, embarrassment, and frustration. A covert test would not require the participants to understand the task to screen their color vision successfully. Therefore, this technique would be useful for the diagnosis of color vision defects in people who might have difficulty understanding and following an explicit testing procedure.

An alternative to explicit responses required in the color vision tests would be to use automatic behavioral responses associated with visual recognition. For example, the galvanic skin response (GSR) is used as one of a number of metrics for determining if a subject is lying during a polygraph test [30]. The GSR is an autonomic system response that is not directly under individual control, but can be indirectly manipulated given that it is related to emotional arousal. Unfortunately, GSR requires physical contact with the participant in order to make a recording, making it less than ideal for covert testing applications.

Another method to diagnose visual disorders without the use of explicit responses of the participant is photoscreening. The American Academy of Pediatrics has proposed the use of photoscreening as an alternative to conventional vision screening tools given that conventional tools can be difficult to administer with individuals that are nonverbal or have developmental delays [31]. Photoscreening uses a camera or video system to obtain images of the pupillary reflexes to diagnose amblyopia and strabismus in young children. Photoscreening cannot currently be used for detecting color vision deficiencies however.

A covert color vision test utilizing eye movements with the preferential looking paradigm has been administered to young children [32]. In this test, various color cards are simultaneous shown to a child from a distance. The administrator covertly observes the child through a peephole in the cards to estimate how reliably a child stares at any given stimulus. The child's behavior is assessed based on the administrator's observation of where the child is looking as determined by considering both head and eye movements. While the procedure is convenient, it's accuracy and reliability depend heavily on the administrator's skill at judging head and eye movements.

An alternative explored in this dissertation is the use of precise eye movement measurement with eye-tracking technology in order to diagnose color vision deficits. This approach minimizes the possibility of human errors possible with the preferential looking technique and eliminates the need for participants to make explicit responses on the test. Under natural viewing conditions, the guidance of eye movements is automatically influenced by visual stimulus properties. For example, it has been repeatedly demonstrated that in well-controlled visual search tasks, display items defined by a single unique feature can attract and even capture attention [33, 34]. Attention and eye movements are strongly

correlated as attention is required to program saccadic eye movements [35]. Furthermore, the eye movements of individuals freely viewing complex natural scenes are also stimulus driven [36, 37]. With addition of eye movement measurement, color vision testing can be indirectly manipulated to evaluate observer's visual performance without requiring explicit responses (i.e., verbal communication) or even the understanding of test instruction from the participants, making the test less sensitive to subterfuge

1.2 Research Objectives

This dissertation focuses on exploring the use of precise eye movement measurement obtained with eye-tracking technology in order to diagnose color vision deficits. The first objective in the study was to extend the Ishihara testing method by using eye movements so that no explicit input is required. This approach facilitates an alternative color vision test that covertly screens for color vision deficiencies. To expand the option of using stimuli other than the Ishihara images, the second objective was to create a method for detecting color vision deficiencies using photographs to make the technique more covert. A technique to select photographs that can be used as stimuli for differentiating viewing behavior of normal and colorblind individuals will be developed. The third objective was to validate the described method by conducting an eye movement experiment in which normal and colorblind individuals were used as participants.

1.3 Dissertation Organization

This dissertation is organized into five chapters. Chapter 1 contains a statement of the problem, research objectives, dissertation organization, and literature review. Chapter 2 presents an experiment that used digital Ishihara images as stimuli to investigate the

differences of eye movements between two groups of participants. Chapter 3 introduces a method to select photographs likely to be perceived differently by normal and colorblind individuals as well as a guideline to identify diagnostic regions of interest from these images. Chapter 4 presents the second eye movement experiments to validate the method. General conclusions and implications of this work are provided in Chapter 5.

1.4 Literature Review

1.4.1 Visual Perception

The processes of visual perception are complex and involve aspects of the physics of light, physiology, and cognitive psychology. In other words, seeing as we know it from everyday experience is the result of interaction between the object, light, the eye, and the brain. The eye acts as the medium for the light to pass through before reaching the decoders that finally transform visual input into electrical signals for the brain. Initially, when light reflects from an object, it travels through the cornea of the eye, and then reaches the pupil that regulates the amount of incoming light by way of contraction of the iris [38]. Then, light reaches the lens where it is focused onto the back of the eye, the retina (see Figure 1.1). The retina consists of light sensitive photoreceptors that transduce the image falling upon it. The retina passively selects which information is vital and which information should be disregarded by sampling the image with a resolution that varies across its surface [39].

The central visual field is sampled by the area of the retina called the fovea [40]. This area provides the highest resolution of visual information but processes only the central five degrees of the visual field [40, 41]. The remaining visual information falls on peripheral areas of the retina providing a lower resolution of visual information. Each area is specialized

for a different purpose. Foveal vision provides a way to acquire highly detailed visual information whereas peripheral vision allows human to perceive the entire scene in order to collect contextual information [41].

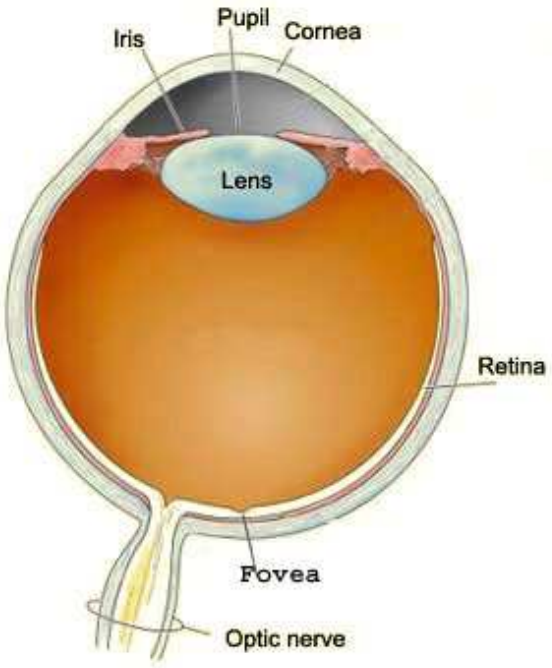


Figure 1.1. Cross-section of the human eye. This image is from the website http://www.wikipedia.org and is permitted to copy, distribute and modify under the terms of GNU Free Documentation License.

1.4.2 Color Vision

The photoreceptors in the retina consist of rods and cones. Rods are responsible for nighttime vision (dim lighting conditions) whereas cones are responsible for daytime vision (bright lighting conditions) [38]. Cones convert light energy into electrical signals that are sent through the visual pathway for the brain to produce color sensation [42]. In a normal human eye, there are three types of cones. Each type is sensitive to different wavelengths of light [43-45]. These types of cones have generally been referred as “red”, “green”, and

"blue" cones. They show approximate maximum absorption in the spectral region of 566, 543, and 445 nm respectively. However, wavelength sensitivities are broad and partially overlapping. For example, the red cones are also somewhat sensitive in the yellow and green parts of the spectrum. Therefore, it is more appropriate to refer these three types of cones as long wave (L) cones, medium wave (M) cones, and short wave (S) cones [40, 46].

Cones are distributed across the retina [47]. However, the probability of absorption of light of a particular wavelength varies across the retina and depends on the cone distribution. The relative mean density of L:M:S cones in the retina of a normal individual is 40:20:1 [48, 49]. The cone distribution can vary among individuals [50]. The fovea contains only cones. Therefore, the fovea provides both color and high resolution visual information. Color vision is limited outside the central 40º of the visual field. Rods are densely packed in the parts of the retina responsible for processing the incoming light from the peripheral visual field. The visual periphery is used primarily for movement detection [51].

1.4.3 Trichromatic Color Theory

Prior to the 19th century, color was believed to be a property of the light rays themselves, not a human response to the wavelength of light. It was Thomas Young who in 1802 introduced the trichromatic theory of color. He hypothesized that human color vision is based upon three types of photoreceptors, sensitive to red, green, and blue primary light. All colors in the visible spectrum can be matched by mixing of three primary colors [52]. This theory was supported by Helmholtz [53] who proposed the existence of three overlapping spectral sensitivity curves corresponding to the three kinds of nervous fibers in the eye. He proposed that light excites these three nervous fibers in various degrees to provide spectral color sensations [38].

Later, Wright [54] and Guild [55] separately conducted experiments to investigate how colors can be matched by mixing the three primaries. Participants observed a circular field that was split in two halves in a darken room. One half was a test field illuminated by a test color while the other half was the mixture field illuminated by three adjustable primary lights (red, green and blue lights). Participants were asked to adjust red, green and blue lights in the mixture field so that they appeared identical to a light in the test field. In the other words, three primaries need to be mixed to match a visible light. However, it was found that not all visible test colors could be matched using this technique. One primary needed to be added to the test color to match with the remaining two primaries. With this procedure, Wright and Guide argued that a whole visible spectrum of colors could be generated by varying the light intensity of three primaries and confirmed the validity of the trichromatic theory.

However, this theory attributes the perception of color to specific light combinations. Therefore, it cannot explain some important color vision phenomena [51]. For example, it cannot describe the simultaneous color contrast phenomenon in which the color appearance of a stimulus is affected by other colors surrounding it. It also cannot account for the metamerism effect when two objects with different spectral properties appear identical in a certain lighting conditions but different in other lighting conditions.

1.4.4 Opponent Color Theory

The opponent color theory originated from visual phenomena that some colors are never seen simultaneously. For example, it can be visualized that a reddish yellow is orange; a reddish blue is purple; a greenish yellow is light green. However, there seems to be no such colors as reddish green or yellowish blue. Proposed by Hering [56], the opponent color

theory proposed that color is processed by bipolar color channels referred to as red-green and blue-yellow channels. The red-green channel responds to either red or green but not both colors at the same time. Similarly, the blue-yellow responds to either blue or yellow. The opponent color theory explains that color appearance depends on the bipolarity of each process. The red and green for example is opponent, they are thus never seen at the same time.

The validity of this theory was verified by DeValois [57] who demonstrated the existence of color opponent neurons by using electrodes connected to the optical nerve of a monkey and argued that the color information is encoded in an opponent fashion. For example, green and red cones oppose each other to produce red-green opponent neurons. The results of this experiment suggested that the three types of receptors exists but the responses of these receptors are converted between the eye and the optical nerve into three other opposed pairs which are red-green, yellow-blue, and white-black opponent cells. Later, Hurvich [51] reported quantitative data in support of the opponent color theory that color perception can be explained by the relative activity in the black-white, red-green channel, and yellow-blue channels. In this experiment, participants were required to adjust the amount of red or green *and* blue or yellow to a mixture field until it turned white. The data were used to create opponent color functions which show the sensitivity of these three opponent channels.

The studies of color vision theories, in summary, have established two facts. First, color vision is trichromatic. The trichromacy indicates that three types of receptors (cones) handle the color matching information. Second, considered as a post-receptor stage, the

opponent color theory described at the neural level how cones interact with each other. Three types of cones oppose each other when providing input to color opponent cells [39].

1.4.5 Description of Color

Color can be described using three perceptual values. These values are hue, lightness, and saturation [51]. In this way, a three dimensional system can be created to represent all perceivable colors. Hue is the attribute of the visual sensation closely associated with the dominant wavelength. For example, a hue of green is psychologically equivalent to a stimulus of 540 nm while a hue of yellow is equivalent to a stimulus of 570 nm. Normally sighted individuals can distinguish about 150-200 hues in the visible spectrum (380 to 780 nm). Lightness is used to refer to the degree of more or less light. Saturation is the attribute of visual sensation used to assess purity of color by varying degree of white content in that hue. Saturation scale can vary from saturated color (no white content at all) to completely unsaturated color (no hue content at all).

Although each color can be characterized by these three values, the relationship between colors is difficult to describe. To provide a convenient way to classify and determine the relationship between colors, the concept of using CIE chromaticity diagram was introduced to specify colors.

1.4.6 CIE Chromaticity Diagram

The CIE chromaticity diagram was proposed with the objective to describe all visible colors mathematically by using three additive imaginary primaries. Under this system, color stimuli can be matched by combination of three imaginary primaries. Imaginary primaries are theoretical positive numbers; therefore, each value does not represent actual visible

colors. Several CIE diagrams have been developed but two of the most commonly used diagrams in color research studies are the CIE XYZ and the CIE LUV.

For CIE XYZ, the amounts of the primaries, given by the upper case letter, are known as the stimulus values [58]. The chromaticity coordinates are given in lower case letters. The relationship of these values is shown in the following equations:

$$x = X/(X+Y+Z) \tag{1.1}$$

$$y = Y/(X+Y+Z) \tag{1.2}$$

$$z = Z/(X+Y+Z) \tag{1.3}$$

and

$$x + y + z = 1. \tag{1.4}$$

A two dimensional plot formed by using two of the chromaticity coordinates to represent colors is called the CIE chromaticity diagram (see Figure 1.2). The horseshoe shaped boundary, referred as the *spectral locus*, represents the pure colors (monochromatic hues in nanometers). The white point falls in the center of the diagram and is indicated by W. All colors which are perceivable by the human eyes fall within the boundary. When three colors are represented by three points on the diagram and lines are projected between the points to form a triangle, all possible colors produced by the addition of those three colors are found within the triangle.

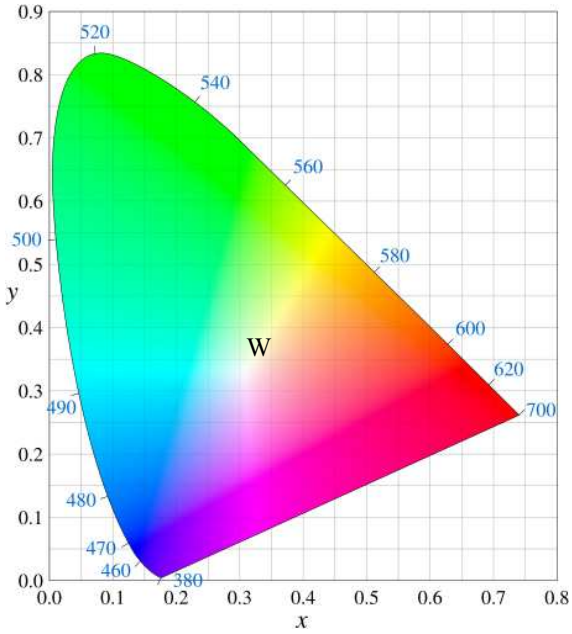


Figure 1.2. CIE XYZ chromaticity diagram. Dominant wavelengths are located on the perimeter and white light (W) is located in the center. This image is from the website http://www.wikipedia.org and is permitted to copy, distribute and modify under the terms of GNU Free Documentation License.

In spite of all the useful properties of the CIE XYZ diagram, it lacks perceptual characteristics. The distance between two points (two colors) is not proportional to the perceived color difference [40]. In order to correct this, the International Commission on Illumination proposed another color space called CIE LUV which better represents uniform color spaces [59]. For example, if color A and B are twice as far as color C and D, then the perceived difference between A and B is about twice the perceived difference between C and D. This diagram is a linear transformation of CIE XYZ, plotting hues on a pair of axes, *u'* and *v'*, with the lightness on the L* scale [59]. The *u'* and *v'* provide chromaticity information whereas the L* encodes the luminance of a given color. The formulae for computing CIE LUV coordinates from the CIE XYZ are given as follows.

$$u' = \frac{4X}{(X+15Y+3Z)} = \frac{4x}{(-2x+12y+3)}$$
$$v' = \frac{9Y}{(X+15Y+3Z)} = \frac{9y}{(-2x+12y+3)} \quad (1.5)$$

$$L^* = 116\,(Y/Y_n)^{1/3} - 16 \quad \text{if } Y/Y_n > 0.008856 \quad (1.6)$$
$$L^* = 903.3\,(Y/Y_n) \quad \text{if } Y/Y_n \le 0.008856$$

where Y_n is luminance value of a standard white.

1.4.7 Color Vision Deficiencies

Normally sighted individuals can perceive a wide range of colors by using all three types of cones (L, M, and S cone) in the retina. They are referred to as trichromats (three colors). Colorblindness is a condition where a person cannot differentiate between certain colors or shades of color because particular cone pigments are missing.

Most commonly, colorblindness is inherited and does not change over time [61]. Approximately 8% of men and 0.5% of women in the North American population are affected by color blindness [19]. Three types of colorblindness are classified according to the cone pigment that is missing or displaced [39].

Monochromats are those individuals who have no color discrimination because two or all three types of normal cone photoreceptor pigments are missing. These individuals are totally colorblind. They only can see light ranging from black to white. A monochromat cannot distinguish objects on the basis of hue or colors alone. However, they may detect differences on the basis of brightness [1].

Dichromats are those individuals who lack one type of cone photopigments. They use only two types of cones photopigments to perceive color. Dichromats include *protanopes* and *deuteranopes*, who lack the L and M photopigments respectively. The lack of L or M

photopigments results in a red-green confusion. Since *protanopes* are less sensitive to red light, reds appear dark brown to them whereas *deuteranopes* see reds as yellowish brown [39]. *Tritanopes* lack the S photopigment, incurring blue-yellow confusion, which is a rare acquired defect.

Anomalous trichromats have three types of photopigments but one of the photopigments has a shifted spectral sensitivity. The displaced photopigment results in an altered color discrimination ability. *Protanomalous, deuteranomalous, tritanomalous trichromats* are those individuals who have abnormal L, M, and S photopigments respectively [1]. The prevalence of deuteranomalous defect is the highest among all types of colorblindness, accounting for about 4.9% of men. On the other hand*, tritanope* and *tritanomalous trichromat* (i.e., blue-yellow defects) are very rare forms of color vision deficiencies (see Table 1.1).

Table 1.1. Classification of colorblindness, occurrence, and color confusion (adapted from Birch [62]).

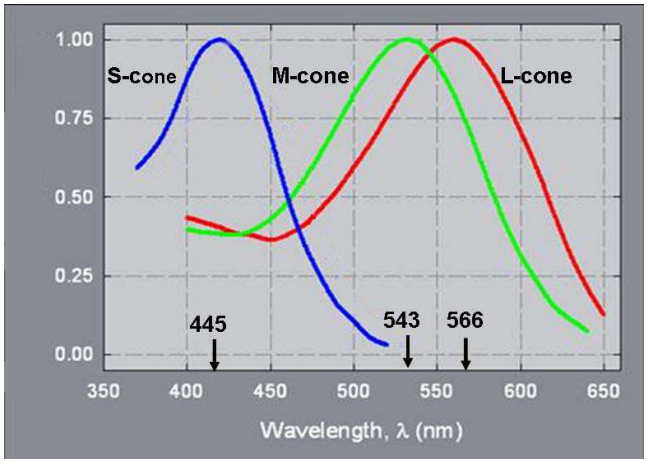
Type	Number of cones available	Classification	Occurrence in North American population (%)		Color confusion
			Men	Women	
Monochromat	One	Monochromat	Very rare	Very rare	No color discrimination
Dichromat	Two	Protanope	0.74	0.02	Bright red/orange/yellow/green, Green/brown, Red/dark gray
		Deuteranope	1.1	0.01	Bright red/orange/yellow/green, Green/brown, Green/dark gray
		Tritanope	0.002	0.001	Violet/yellow-green, Yellow/white, Blue/black
Anomalous trichromat	Three	Protanomalous	1	0.02	Red/orange/yellow/green, Red/dark gray/brown
		Deuteranomalous	4.9	0.38	Red/orange/yellow/green Green/dark gray/brown
		Tritanomalous	Rare	Rare	Violet/yellow-green, Yellow/white

As seen in Figure 1.3b for anomalous trichromat, the wavelength of peak sensitivity for L cone is different in a protanomalous and a normal individual (555 nm vs. 566 nm), resulting in a reduction in sensitivity for red-green shades. For a dichromat as seen in Figure 1.3c, the absence of the L photopigments means that the spectral sensitivity is based on the absorption capability of just the S and M cones. This leads to a huge reduction in sensitivity to longer wavelengths. In this case, bright red will be perceived as dull brown. Therefore, protanopes, deuteranopes, protanomalous, and deuteranomalous trichormats have a similar reduction in red-green discrimination ability, varied only in magnitude.

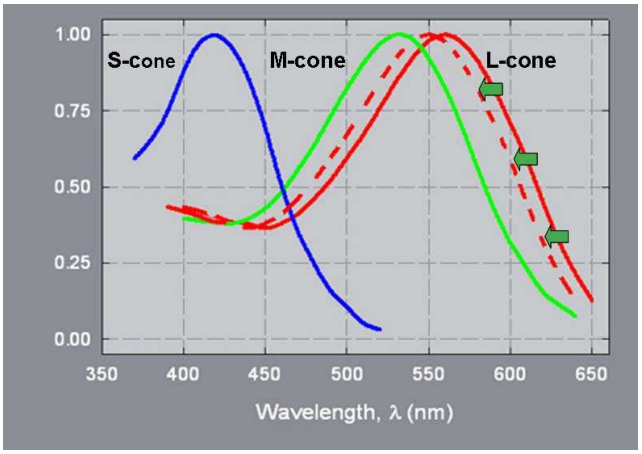
1.4.8 Red-Green Colorblindness

Red-green colorblindness is generally a collective term referred to those who have problem with discriminating the range of red-orange-yellow-green. The term includes both dichromatic and anomalous trichromatic groups. People with severe red-green colorblindness (dichromats and severe anomalous trichromats) typically share a problem of distinguishing saturated reds from greens while those with mild red-green colorblindness (mild anomalous trichromats) are confused with de-saturated red-greens. In this dissertation, the term "colorblindness" was used to refer to red-green color deficiencies including protanopia, deuteranopia, protanomalous trichromacy, and deuteranomalous trichormacy. As can be seen in Table1.1, these conditions make up the vast majority of color vision deficiencies in the population.

(a)



(b)



(c)

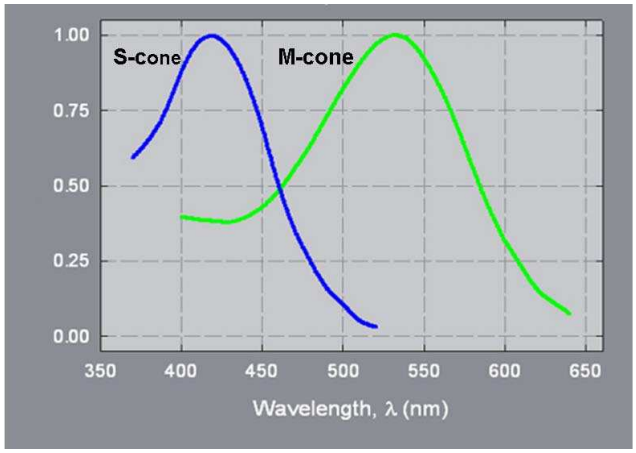
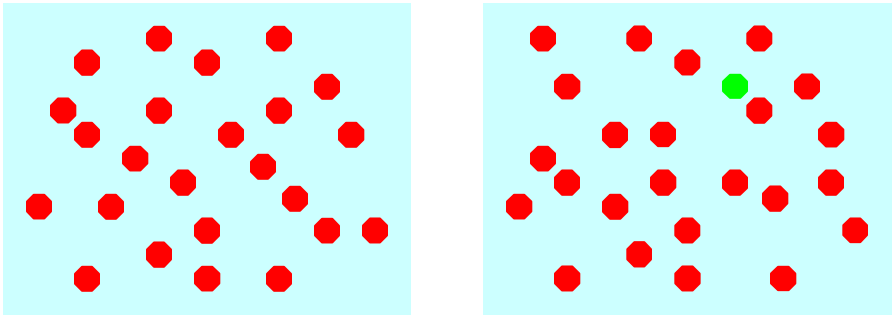


Figure 1.3. Sensitivity of human cone as a function wavelength. (a) Normal observers whose eyes contain three types of cones (b) Anomalous trichromats whose L cones abnormally shift (i.e., protanomalous) (c) Dichromats whose L cones are absent (i.e., protanope). Both groups have trouble distinguishing reds and greens. Adapted from [63] with permission of Bruno Dubuc.

1.4.9 Visual Attention

To better understand what draws observer's attention, and how they divert their attention to a particular location when a stimulus is presented, the mechanisms of visual attention are explored in this section. The main concept of visual attention involves selection. This means that human visual system selects only some parts, not all, of visual input to process [36, 64, 65].

Two distinct mechanisms of visual attention are bottom-up and top-down mechanisms of selection. The bottom-up selection is automatic and stimulus driven. Stimulus properties of an object can attract attention in a compulsory manner [64, 66, 67]. In the other words, visual attention is directed to a particular object in a scene due to obvious feature of colors, shape, brightness, motion, or orientation of the object. For example, a green circle is easily noticed at a distance among red circles (see Figure 1.4). This target exhibits a pop-out effect due to its distinguishable color.



(a) (b)

Figure 1.4. An example of a pop-out effect. (a) target is absent in red circles (b) target is detected preattentively due to its green color.

On the other hand, the top-down selection is task dependent or goal oriented [64, 66, 67]. The observer's eyes are focused only on the locations related to that task. This type of

visual attention is usually slower and independent of stimulus features. For example, looking for a flight schedule on an airport monitor screen, observers will seek for the flight number information, overlooking nearby colorful airline logos.

The concept of selectivity can be usually used to explain such phenomena as the ability to filter out unimportant information [68]. When a target of attention is different from the rest of the surroundings, the rest of the information has little effect on visual attention. In this case, the target quickly captures attention via bottom-up mechanisms. In some cases, even given top down mechanisms where an observer's attentional guidance is goal-directed, the bottom-up stimulus factors still play an important role the guidance of visual attention [68].

However, the argument of which mechanism (top-down vs. bottom-up) has more influence on attention is still unresolved. Bacon and Egeth [64] proposed a concept of search modes that postulates that when observers utilize a "*singleton detection mode*", the location with the largest feature contrast will attract attention. The notion of singleton means a visual attribute (e.g., color, orientation, or intensity) that differ from background. In this mode, relevant or irrelevant singletons can capture attention, depending on which location provides the greatest feature contrast. On the other hand, if observers are engaged in a "*feature search mode*", only the location that matches the task-related feature is getting attention. In this mode, irrelevant singletons will not distract visual search for a singleton target under top-down influence. The degree to which of these two modes assumes priority is described by the target-distractors similarity [69]. If the target is not similar to the distrators, singleton detection mode tends to be used. If the target and distractors are highly similar, feature search

mode will likely be engaged. These findings indicate that visual attentional priority depends on task requirements [70, 71].

Nevertheless, the singleton detection mode is assumed to be a default visual search behavior [64, 70]. A similar concept was used by Koch to develop a saliency model of bottom-up attention. This model predicts the areas where normal observers will attend in photographs [72]. In this model, a highly salient feature will draw attention [73]. Parkhurst et al. [36] investigated the validity of this saliency model and found a correlation between the stimulus salience predicted by the model and observed human eye fixations under natural viewing conditions. The saliency model has become a dominant theoretical view as it has been cited by several literatures [36, 71, 74] and studies of eye movements in both cognitive psychology and computer vision fields.

1.4.10 Computational Model of Saliency

Described in this section is a computational model of saliency that takes as input an image of a scene and produces as output a prediction of locations in the scene where observers will attend. The saliency model is based on the findings that salience is determined rapidly by the low-level visual system [75]. Low-level features include color, intensity and orientation, each competes to be the highest salient feature. The most salient features will be represented as locations where attention should be directed [72, 75]. The processing in the saliency model is as follows.

First, the input image is decomposed in parallel in three feature pathways, which are color, intensity, and orientation (see Figure 1.5). Each pathway is filtered at eight spatial scales using a Gaussian pyramid scheme. The scale ranges of the pyramid ranges from 1:1 to 1:256, reducing the image resolution by a factor of 2 [36]. This enables assessment of salient

regions at different sizes. For the color pathway, it generates pyramids of red, green, blue, and yellow. For the intensity pathway, it generates a pyramid of average luminance from all color components. Finally, for the orientation pathway, it provides 4 pyramids corresponding to 0˚, 45˚, 90˚, and 135˚. These orientations are used because they simulate the receptive field structure present in the human visual system [73].

Second, for each feature within each pyramid, the differences between a fine and a coarse scale are calculated to simulate *the center-surround mechanism* of the neural receptive field of the human vision system [73]. If light falls on the center of the receptive field, it has a stimulating effect. However, if light falls on the surround, it will have an inhibitory effect [42]. The calculated differences between pixels corresponding to the center of the receptive field in a fine scale and pixels corresponding to the surrounding of the receptive field in a coarser scale generate opponent features that centrally stand out from their surroundings. From this process, the color channel provides two center-surround pyramids as representatives of the opponent color system: 1) center-surround differences in red and green and 2) center-surround differences in blue and yellow. The intensity channel provides one pyramid by determining the center-surround difference of intensity within the pyramid. The orientation channel provides four pyramids, each is calculated by the center-surround difference of intensity within its pyramid.

Third, each center-surround map is normalized by the maximum saliency value obtained at that center-surround pyramid. Pyramids within each channel are then linearly summed to create one pyramid (i.e., several feature maps on different scales). Next, these maps of single channel are summed up into a feature map to a resolution of 40x30, each for color, intensity, and orientation. The feature maps are later normalized to replicate

competition for salience within each feature map (i.e., within-feature competition) across the scales and finally are combined to create the saliency map [36].

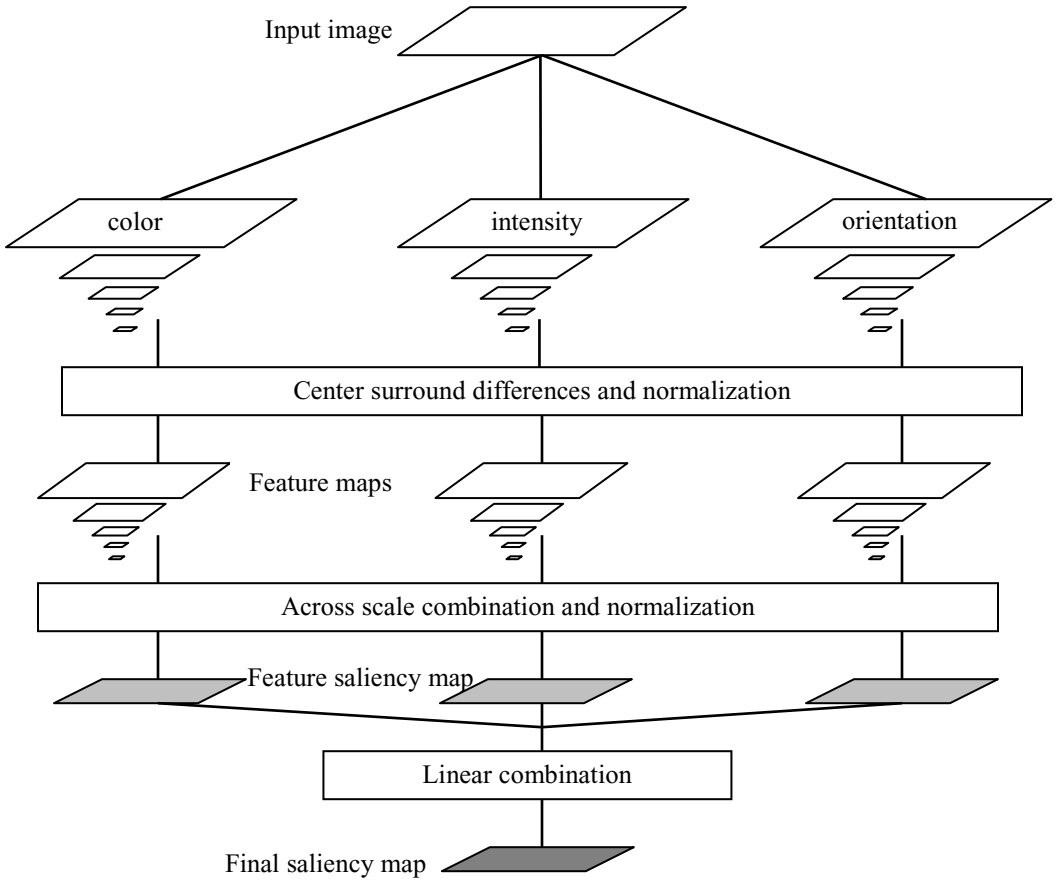


Figure 1.5. Saliency model diagram [75]. Original image is decomposed into 3 features with linear filtering at 8 spatial scales. The center-surround difference mechanism is implemented by subtracting center at a fine scale with the surround at a coarse scale. Maps of single features are then summed up across scales to create a single feature saliency map (color, intensity, orientation). Three feature saliency maps are then combined and normalized to create the final saliency map.

1.4.11 Eye Movement and Visual Attention

Several studies have extensively investigated the relationship between visual attention and eye movements and found that eye movements can be used to explain visual search behaviors [76-78]. As the activities of the visual muscular system, eye movements can

provide insight on the human attentional processes [79, 80]. Although there are several types of eye movements, of most interest to cognitive science studies are saccades and fixations. Saccades are rapid eye movements that move the point of gaze from one target to the next. Saccades can be made as quickly as 500 degrees of visual angle per second [81]. Humans therefore do not obtain new information during a saccade given the presence of blur [82]. Typically, researchers measure the movement between each saccade in terms of saccade length or degree of visual angle. Their pattern of locations is called the scanpath. A fixation is a relatively steady period of gaze which lasts about 200-300 milliseconds at a particular location [81]. It is during this moment that high-resolution visual information can be collected. Researchers generally measure a fixation in terms of its location, dwelling time, and frequency on targets as a metric to assess visual attention.

As partly discussed in the previous section, the visual field can be separated into foveal, parafoveal, and peripheral based on visual acuity [39]. Although visual acuity is sharp in the fovea, it is poor in the parafovea (about five degrees from either side of fixation), and it is the poorest in the periphery (area beyond the parafovea). Therefore, humans need to execute saccades to fixate on different parts of the stimulus in order to see them clearly.

There are also other types of eye movements such as pursuit, vergence, and vestibular eye movements [81]. Pursuit eye movements maintain a point of gaze on a moving target. During pursuit, visual details of the object can be made out but the background is blurred. Vergence occurs when eyes move inward to fixate on an object close to the viewer. Vestibular eye movements occur when the eyes rotate to compensate for head or body movements. Although pursuit, vergence, and vestibular are essential to human perception,

saccadic eye movements and fixations are of primary concern in the research presented in this thesis.

1.4.12 Eye Tracking

Eye tracking can be used to determine locations of gaze. Eye tracking has been gaining in popularity over the past decade as a method to study cognitive behavior in such domains as image viewing, reading, and driving situations [76]. There are a variety of eye movement measures that can be used depending on the purpose of the study. Some of the most commonly used measures include fixation location, fixation duration, number of fixations, saccade length, and scanpath [78].

Eye tracking techniques can be categorized into analog and video based methods [83]. The analog methods use several techniques. For example, electro-oculography measures eye movements by placing electrodes near the eyes. When the eye muscles move, changes in electrical activity can be detected. Contact-lens based eye coil systems tracks eye movements via affixed objects in contact lens. This contact lens contains an affixed mechanical lever or magnetic coils used to detect movement. These techniques however have not been prevalently used as compared to the video based eye tracking methods.

Video based methods generally utilize visible or infrared imaging, and are either remotely located or head-mounted. Visible spectrum imaging captures ambient light reflected from the eye. This technique tracks the contour between the iris and the sclera known as the limbus. The disadvantage of this method is it depends on the uncontrolled ambient light of the source. Infrared imaging uses a uniform infrared light to illuminate the eye. This technique tracks the contour of the pupil in the image. Tracking the pupil contour is

preferable because of its smaller size, shaper contour as compared to the limbus, and greater visibility under infrared illumination.

Infrared imaging methods commonly use a bright pupil technique. The retina of the eye is very reflective to infra-red light, enabling a bright-pupil image of the eye to be detected on a sensor when the eye is illuminated with a light source in the same or a close optical path as the camera. The specular reflection of the illuminating light off the front surface of the cornea is known as the corneal reflection. The corneal reflection and the bright-lit pupil together are used to determine the point of gaze [84]. On the other hand, the dark-pupil technique illuminates the eye with an off-axis light source so that the pupil appears as the darkest spot in the image, while the sclera, iris, and eye lids reflect the illumination. The corneal reflection can also be detected with this technique. The corneal reflection and the dark spot pupil are used to determine the point of gaze.

Eye tracking systems are either remotely located or head-mounted. The remote system is an unobtrusive system in which the participant's eye movements can be tracked via a pan/tilt camera. This camera contains an eye illuminator and moves automatically to track the participant's head. The eye is illuminated by the beam from near infrared LED's on the pan/tilt camera. The remote system uses a control unit to process the eye image and extract the location of pupil and specular reflection of the light source. The unit computes both pupil diameter and line of gaze. The head-mounted system uses the head-mounted optic module to limit the tracking error from head movements. It usually has greater accuracy than the remote system because it measures the line of gaze with respect to the head.

To accurately track eye movements, a calibration process of setting up the eye tracker to associate screen positions with pupil and corneal reflection positions is needed. The

purpose of calibration is to provide data that will allow the eye tracker processor to account for individual differences [85]. If calibration is not performed correctly, eye tracking errors will be generated. This process requires two necessary steps. First, the eye tracker defines the nine calibration points with respect to the monitor image. Then it records the participant's fixation for each point to compare with the calibration target. The relation between these two will be used to adjust to individual differences [86]. Generally, the nine point calibration grid is used as calibration target points. These points cover about 80 percent of the scene monitor screen area and are separated by 15-20 degrees visual angle horizontally, and 10-15 degrees vertically. The accuracy of the tracking is assessed by plotting the fixation coordinates with respect to the actual calibration image.

CHAPTER 2. EXTENDING THE ISHIHARA TEST USING EYE MOVEMENTS

The interpretations of color vision status of the Ishihara test primarily depends on verbal input for which it is possible to learn the correct response to pass the test [12]. In this typical test paradigm, the observer takes the visual information as input and provides verbal information as output, thereby providing a leeway to subterfuge. The motivation for conducting this experiment is to explore an extended paradigm that takes eye movements as a measure and directly uses this measure for interpreting color vision status as an effective means at limiting subterfuge.

When administering the Ishihara color vision test, those with defective color vision will not perceive the same numerals as those with normal color vision [87]. The colors of the numeral and background are confused by the colorblind observer. In this task, colorblind observers cannot find the visual cues that are important for reading the numerals. Figure 2.1 demonstrates the Ishihara images and a color simulation that shows the image as would be seen by colorblind individuals.

Given that stimulus properties influence the guidance of eye movements, and that normal individuals and individuals with color deficiencies sense colors differently, it was predicted that there will be observable differences in the eye movements of these groups even when identical stimuli are observed. Such differences could provide a metric for a covert color vision test. To test this possibility, a study was designed where the eye movements of individuals with red-green colorblindness and individuals with normal color vision were measured.

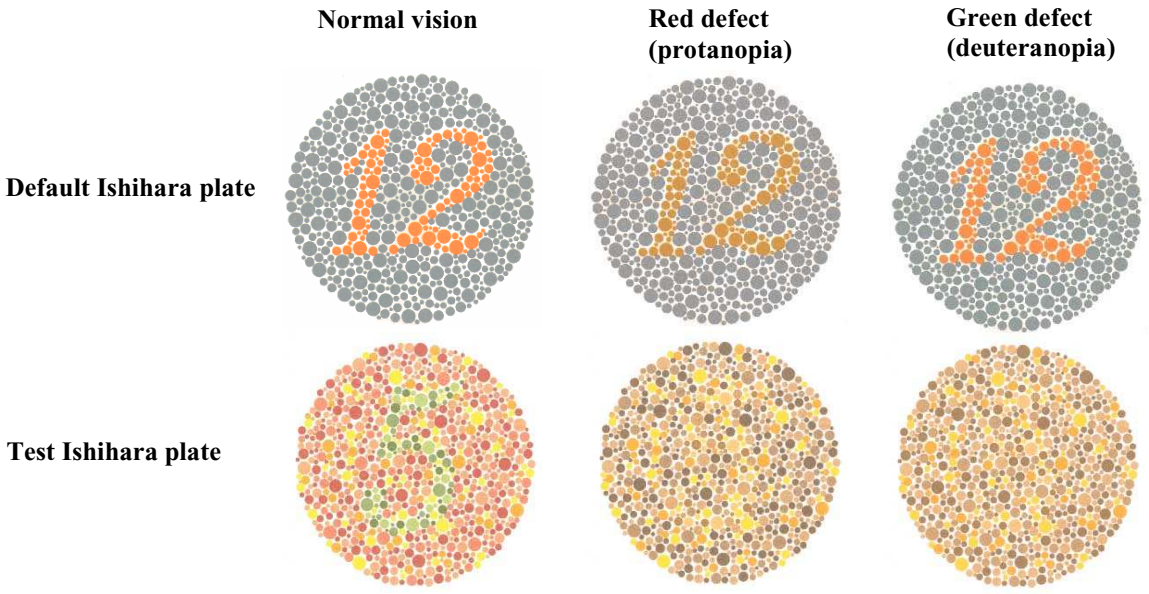


Figure 2.1. Ishihara images as perceived by colorblind individuals. On the first row of images, both normal and colorblind observers can perceive number 12. On the second row, normal observers can see number "5" while observers with red and green defects cannot. The images in the second row were simulated as they would be seen by red and green colorblind observers using Meyer and Greenberg's dichromatic reproduction [88].

Eye movements were recorded in the first block of an experiment where participants freely viewed a series of 15 Ishihara color vision test images. Eye movements were also recorded in the second block of the experiment where participants freely viewed the images again, but were required to identify the numerals in the images, as is required in the Ishihara test. The use of these two tasks allowed investigating not only if eye movements can be used to diagnose colorblindness during the standard Ishihara test, but also outside of an explicit color vision test, during free viewing of the Ishihara images.

2.1 Experimental Methods

2.1.1 Participants

Twenty male students from Iowa State University participated in the experiment. A statement of informed human consent was obtained from all the participants. Each participant was paid $10 for participation. Ten participants were selected on the basis of having been clinically diagnosed as colorblind at some point in their life. The term colorblind was used to describe this group of individuals. Note however that the type of color deficiency (e.g., dichromacy or anomalous trichromacy) for each participant has not been determined. The remaining ten participants were selected as a comparison control group based on a self-report of normal color vision. To confirm their color vision status, all participants were administered the computerized Ishihara test [89] during the second block of the experiment. The interpretations of color vision status from the computerized Ishihara test results were all in accordance with the self-report of the participants.

2.1.2 Stimuli

The digital versions of images in the Ishihara test were acquired from a digital image archive dedicated to visually impaired people (http://www.agingeye.com) and used as stimuli. The fifteen images from the 24-plate edition were used. All of the selected plates displayed the numerals that can be seen by a normal observer but seen differently or cannot be seen by most observers with red-green colorblindness according to the Ishihara test manual [87]. To confirm that the acquired images provide the designed effects where only participants with normal color vision could discern numerals correctly, Meyer & Greenberg's color defective simulation [88] was implemented to synthesize how they were perceived by the colorblind individual. The simulation transformed the hues distinct only to normally

sighted individuals into the hues perceived by color defectives. The results of simulation verified that the correct numerals on each image were disguised or could not be discerned either by the outline or by the tonal changes. Representative images are shown in Figure 2.2.

(a) Normal trichromat	(b) Red-green defect simulation		(c) Numeral displayed
	Protanope	Deuteranope	

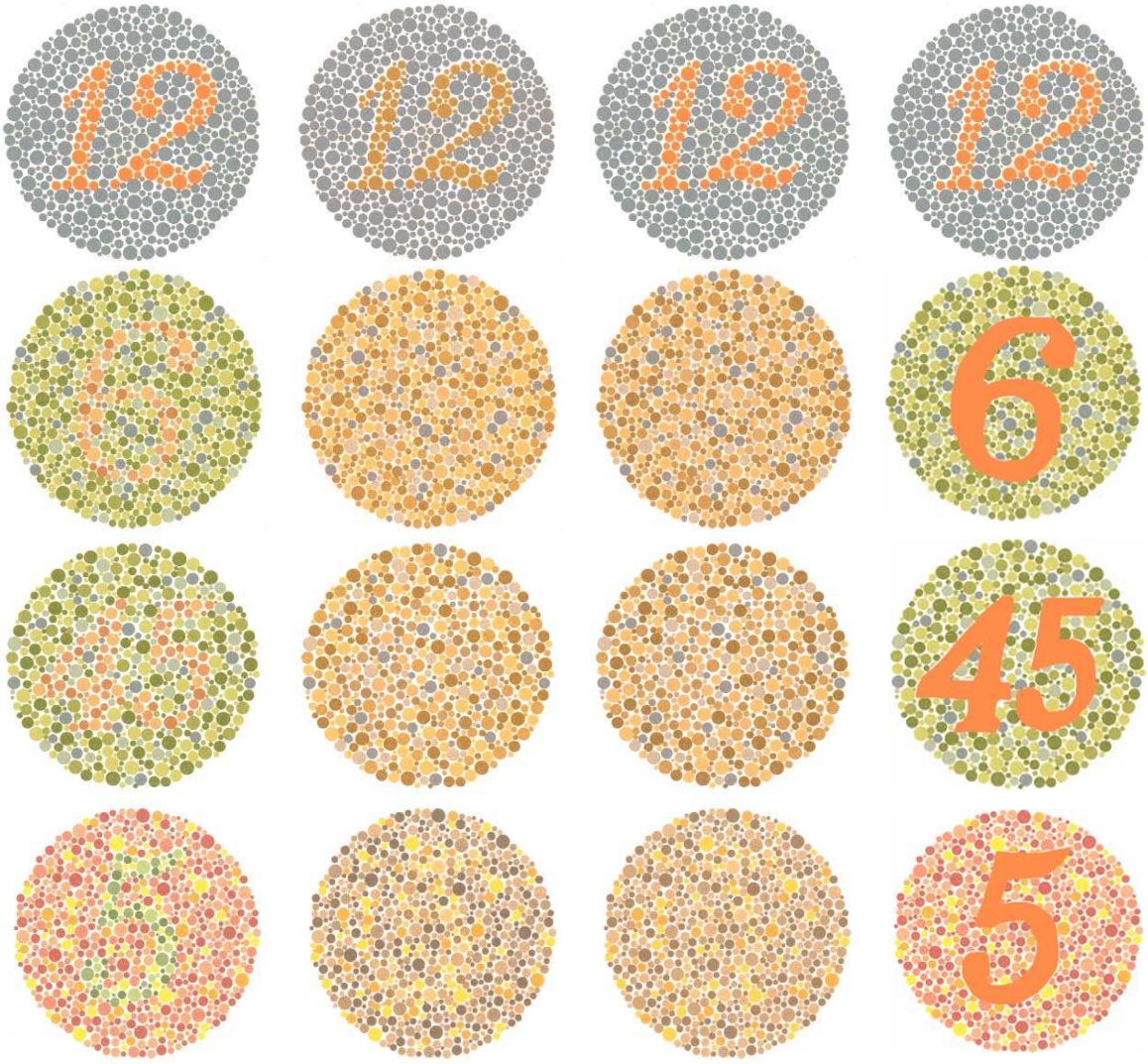


Figure 2.2. (a) Examples of the Ishihara images used in the experiment. (b) Red-green defect simulation of Ishihara images as would be seen by colorblind observers. (c) Numerals on each image as would be seen by normal observers.

2.1.3 Apparatus

The first image of the series is the only image designed to present a numeral that both normal and colorblind participants can report (Figure 2.2; first row). The image size was digitally adjusted to display full screen at a resolution of 1024 x 768 pixels in the 32-bit color mode. Each Ishihara image was presented on a 19-inch monitor (SyncMaster, Samsung, CA). The center of the display was aligned at the participant's eye level at a distance of 75 centimeters. This provided an optical image subtending 27˚ horizontally and 24˚ vertically that facilitate the tracking of eye movements.

The established procedure for utilizing the Ishihara color vision test on a computer screen was adopted[89, 90]. This procedure has been validated despite the differences between the spectral emission of the monitor and the reflected daylight of the Ishihara plates. The computerized method is capable of detecting red-green defects for occupational health purposes, and has a sensitivity comparable to the conventional printed Ishihara test [89-91].

A colorimeter (ColorCAL, Cambridge Research Systems, Kent, U.K.) was used to calibrate the monitor by adjusting the values of luminance, CIELUV-coordinates of D65 white, red, green, and blue colors of the monitor as close as possible to the suggested values[91, 92] (see Table 2.1). Five positions on the monitor (center, top right, top left, bottom right, and bottom left) were repeatedly calibrated and measured to obtain average values for each color. The screen calibration process was performed as an iterative measurement to obtain accurate screen specification appropriate for a color display experiment [93].

Table 2.1. Suggested values of luminance, CIELUV-coordinates of the RGB primary colors of the monitor recommended by Hoffmann & Menozzi (1998) and measured values used in the study.

Color	Luminance (cd/m^2)		u'		v'	
	Suggested	**Achieved**	**Suggested**	**Achieved**	**Suggested**	**Achieved**
Red	25.9	26.1	0.428	0.437	0.527	0.535
Green	76.2	78.3	0.122	0.112	0.562	0.547
Blue	9.14	9.65	0.168	0.155	0.189	0.175
D65 White	101.7	102.8	0.198	0.186	0.475	0.446

An eye tracker (ASL 504, Applied Science Laboratories, MA) was used to remotely track eye movements by using a pan/tilt camera located between the computer screen and the participant. The eye movements were recorded at a sampling rate of 60Hz. Eye tracking software (ET6, Applied Science Laboratories, MA) was customized and used to record eye movement data files while the software (EyeNal 2.18, Applied Science Laboratories, MA) was used to extract fixation locations and fixation durations. According to the ASL instructional documentation, a fixation is defined as the mean coordinates of gaze position lasting at least 100 milliseconds during which gaze point remains within 1˚ of visual angle. A chin rest was used to limit head movements, for the purpose of eye tracking. Figure 2.3 shows the experimental setup. The participants sat in front of the computer screen and fixed their head with the chin rest to minimize eye tracking error.

Figure 2.3. Eye tracking experimental setup.

2.1.4 Procedures

The experiment contained two ten-minute blocks separated by a five minute break. For both blocks, participants were seated in front of the computer screen with their head resting on a chin rest. At the beginning and the end of each block of the experiment, participants were required to sequentially fixate nine fixation crosses on a full-screen grid. These measurements were required in order to calibrate the eye tracker and determine the average accuracy of the eye movement measurements. Since eye tracker system suffers from non-linearity and noise (e.g., a compression of signals as fixation deviates from the axis of the camera), the calibration method provides a means of linearizing the system, thereby providing a metric for associating fixation with screen positions [36]. The average eye tracking error at the beginning of the block was 0.93˚ while average error at the end was 1.05˚. There were no significant differences in average error between groups with the exception that one participant repeatedly failed to follow instructions, and was thus omitted from all data analyses due to a complete failure of the system to track eye movements.

During the experiment, each of the Ishihara images was presented for a period of ten seconds. Between each image presentation, a black dot at the center of the screen was presented for one second to be fixated prior to each trial. In the first block of the experiment, participants were instructed to freely look around at the images. In the second block, the same set of the Ishihara images were presented again. Participants were asked to read and identify the number that they perceived on each Ishihara image. To prevent head movements from affecting the eye movement recordings, participants were required to wait until each image disappeared, and then, respond verbally. Their answers were recorded by the experimenter.

2.2 Data Analysis

To study the locations where the participants tended to fixate, each Ishihara image was divided into two regions: the foreground and the background. The foreground (f) was defined as the area where the pattern of colored dots formed one or more numerals. The background (b) was defined as the remaining area of the Ishihara image that contained colored dots. To quantify the proportion of fixations that a participant made on the foreground, the ratio (f*) of the number of fixations on the foreground (f) relative to those on the total area (f + b) was calculated for each image as:

$$f^* = \frac{f}{(f+b)} \tag{2.1}$$

To estimate the proportion of fixations on the foreground expected by chance factors alone (f_c*), the proportion of pixels on the foreground area to pixels in the total area was calculated (see Figure 2.4). The average f_c* across all images was 0.188.

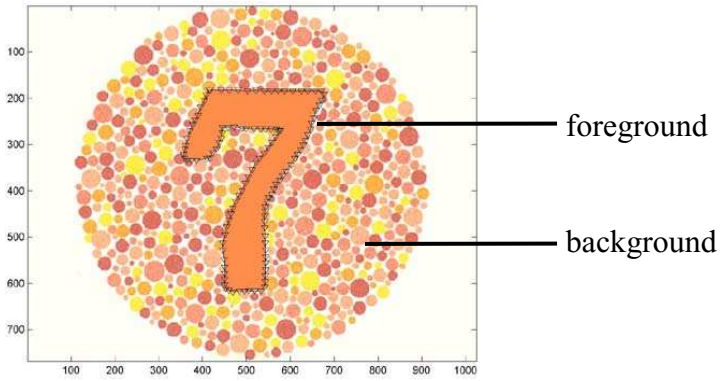


Figure 2.4. Proportion of fixations on the foreground expected by chance factors alone (f_c*) was estimated from the average proportions of foreground in the total area across 15 images.

Given that participants naturally tend to fixate in central regions of the images (see Figure 2.5), a more conservative method of estimating chance performance that accounts for this central bias can be had by using an image shuffling technique [36, 37]. This technique randomly shuffles participants' observed scan paths onto different images prior to calculating f* in order to disassociate fixation locations and image content while retaining the distributional properties of the fixations (i.e., centrally biased). If this more conservative chance performance baseline and the f* differ, it suggests that participants' fixations are indeed guided by image properties, and are not due solely to the chance of fixating centrally.

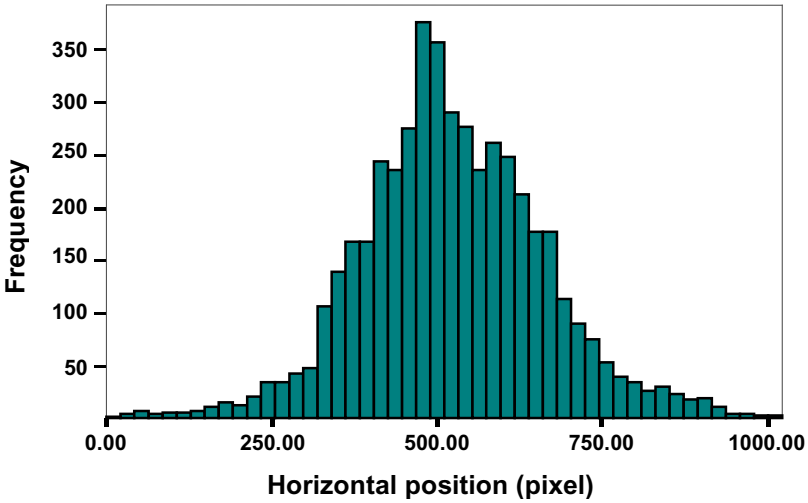


Figure 2.5. Frequency distribution of all horizontal fixation positions on the display screen. Note that the highest frequency is in center of the screen.

Practically, the image-shuffled estimate of the proportion of fixations on the foreground expected by chance factor alone (f_s*) can be calculated by overlaying each participant's observed fixation locations for any given image onto all other images. Thus, a total of 13 f_s* values for each of 14 images were calculated for each block of each participant. Note that the first trial is not included in the analysis because it is a catch trial that portrays a numeral detectable by both normal and colorblind individuals. The average f_s* was calculated separately for normal and colorblind participants given that the distributions of fixations significantly differed between these two groups (see the Results section). The average f_s* across the normal participants was 0.430 while the average f_s* across the colorblind participants was 0.342. Note that these values are much higher and thus more conservative estimates of chance performance than f_c* = 0.188.

Three other eye-movement measures were also analyzed, the total number of fixations per trial, the average fixation duration, and the average saccade length. Each dependent measure was analyzed using a repeated-measures ANOVA with task (free-view

vs. read) as a within-participant factor and vision (normal vs. colorblind) as a between-participant factor.

2.3 Results

Numeral Identification. The criteria to assess a color vision defect is a minimum of four incorrect readings for images 2-15. All numerals were identified correctly by all 10 normally sighted participants. As expected, performance on the numeral identification task for images 2-15 was poor for the colorblind group, with an average accuracy of just 17 percent. Also as expected, all participants correctly identified the numeral in the first image, a default plate which was designed to be perceived as "12" by both groups. The summary of the reading results is shown in Table 2.2.

Proportion of Fixations on the Foreground (f*). As shown in Figure 2.6, the average f* of the normal group was 0.584 (SD = 0.098) when free viewing and 0.625 (SD = 0.082) when reading whereas for the colorblind group it was 0.412 (SD = 0.105) when free viewing and 0.421 (SD = 0.106) when reading. Significantly fewer fixations were made on the numerals by the colorblind group as compared to the normal controls ($F(1, 17) = 390.57$, $p < .05$). A significant interaction was observed between task and vision ($F(1, 17) = 9.72$, $p < .05$). A post-hoc, pairwise comparison using the Bonferroni adjustment ($\alpha = 0.05$) indicated a significant increase in the proportion of fixations on the foreground for the normal group when reading as compared to free viewing. No significant difference across task was observed for colorblind group. Finally, the marginal means of both groups were significantly higher than that expected by chance factors alone, considering either f_c* or f_s*.

Table 2.2. Results of Ishihara reading.

Ishihara image no.	Number of normal participants who read correctly (out of 10)	Number of colorblind participants who read correctly (out of 9)
1	10	9
2	10	0
3	10	1
4	10	0
5	10	4
6	10	1
7	10	2
8	10	0
9	10	0
10	10	0
11	10	2
12	10	0
13	10	0
14	10	5
15	10	6
Number of correct reading (images 2 – 14)	140	21
Percent of correct reading (images 2 – 14)	100%	17%

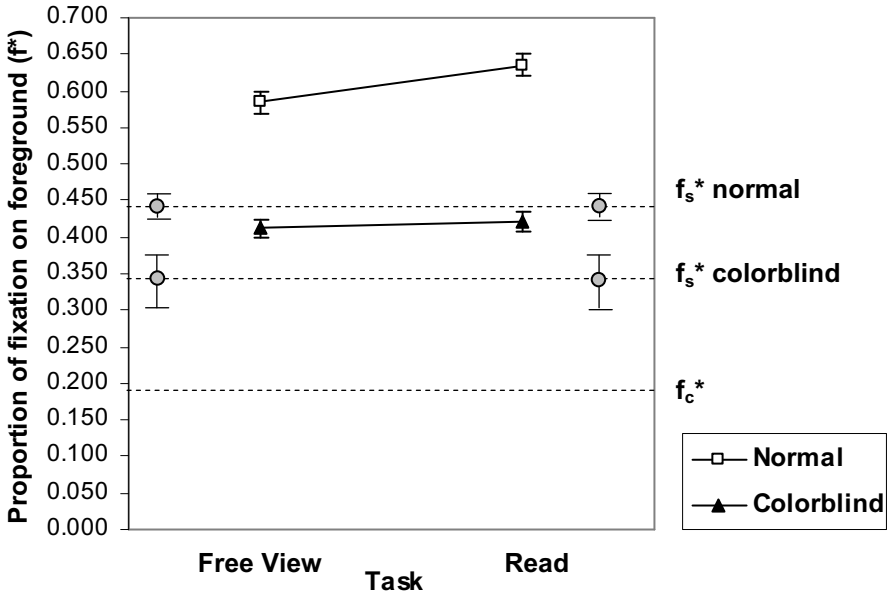


Figure 2.6. The proportion of fixations on the foreground was significantly lower for colorblind participants than normal controls. The proportion of fixations on the foreground expected by chance factors alone are shown as a dashed lines with error bars indicating plus and minus one standard error of the mean.

Fixation Duration. The average fixation duration of the normal group was 0.36 seconds (SD = 0.26 seconds) when free viewing and 0.33 seconds (SD = 0.24 seconds) when reading whereas for the colorblind group it was 0.30 seconds (SD = 0.28 seconds) when free viewing and 0.25 seconds (SD = 0.27 seconds) when reading. The average fixation duration for the colorblind group was significantly less than that for the normal controls ($F(1, 17) = 28.10$, $p < .05$). The average fixation duration in the reading task was significantly less than in the free viewing task ($F(1, 17) = 17.24$, $p < .05$).

Number of Fixations. The average number of fixations of the normal group was 28.17 (SD = 7.35) when free viewing and 31.53 (SD = 9.03) when reading whereas for the colorblind group it was 35.07 (SD = 11.08) when free viewing and 37.24 (SD = 13.29) when reading. The average number of fixations for the colorblind group was significantly higher

than that for the normal controls ($F(1, 17) = 37.56$, $p < .05$). The average number of fixations in the reading task was significantly higher than in the free viewing task ($F(1, 17) = 17.41$, $p < .05$).

Saccade Length. The average saccade length of the normal group was 2.65 degrees (SD = 1.26 degrees) when free viewing and 2.60 degrees (SD = 1.02 degrees) when reading whereas for the colorblind group it was 3.23 degrees (SD = 1.44 degrees) when free viewing and 3.19 degrees (SD = 1.32 degrees) when reading. The average saccade length for the colorblind group was significantly higher than that for the normal controls ($F(1, 17) = 83.58$, $p < .05$).

2.4 Discussion

The goal of this study was to test the prediction that there would be observable differences in the eye movements of normal and colorblind individuals even when identical stimuli are observed by both groups. The proportion of fixations on the numerals, the average fixation duration, the total number of fixations, and the average saccade length were measured. Consistent with this prediction, there were significant differences for each of these measures.

The proportion of fixations made on the numerals by the colorblind group was significantly lower than that made by the normal controls. This result, together with the fact that numeral identification performance was much lower for the colorblind group, supports the assumption that these participants have difficulty sensing particular colors and thus also have difficulty perceiving the numerals. Figure 2.7 demonstrates an example of overlaid fixations made by a normal and a colorblind individual.

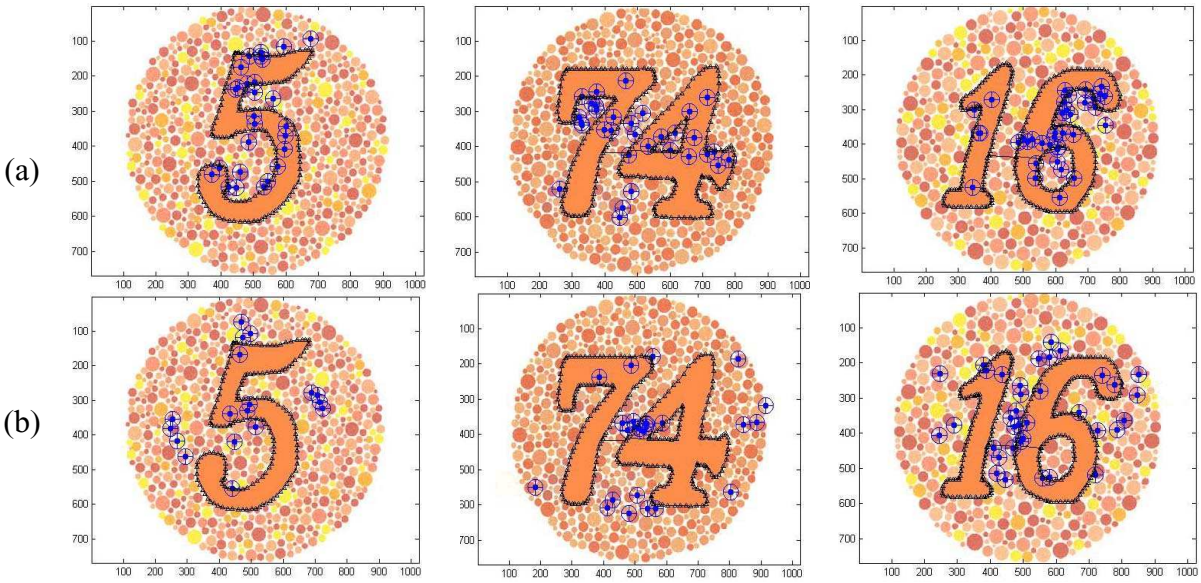


Figure 2.7. Example of fixation patterns (a) from a normal participant (b) from a colorblind participant. Note that fewer fixations were made on the numerals by a colorblind as compared to a normal individual.

Interestingly, the proportion of fixations made on the numerals by the colorblind group was significantly higher than either of the chance estimates. To investigate this result more, the proportion of colorblind participants who answered correctly on any given image as a function of the average proportion of fixations on the numeral for that image was plotted. As can be seen in Figure 2.8, identification performance was indistinguishable from chance levels for the large majority of the images. However, for some images, a higher proportion of fixations on the numeral was associated with better identification performance. This result is consistent with anecdotal reports from participants that for some images, a small portion of the numeral contour was visible and that this visual information aided in guessing the numeral's identity. This implies that some of our colorblind participants could have a mild form of color deficiency. In particular, it is reasonable to suspect that some of the participants are anomalous trichromats.

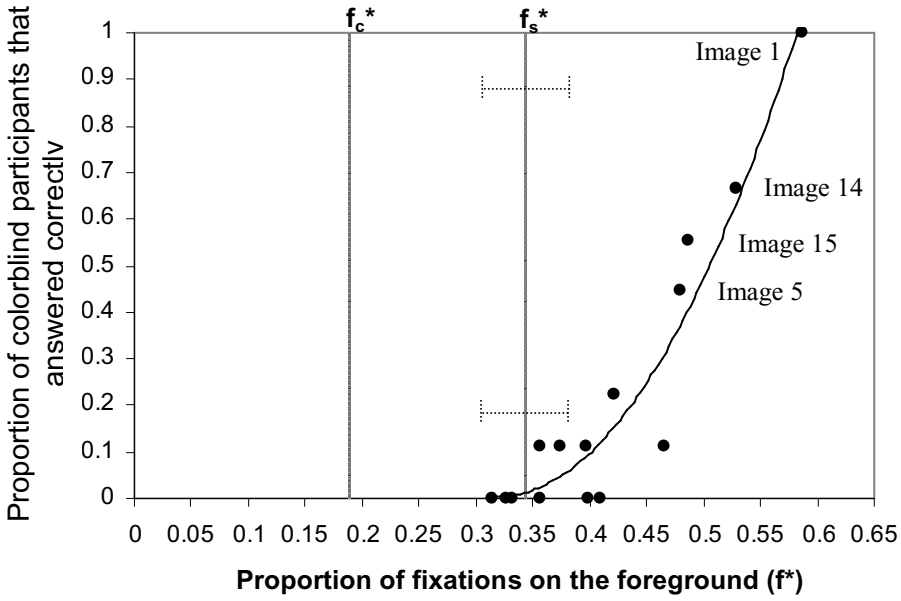


Figure 2.8. The proportion of colorblind participants that answered correctly was positively related to the proportion of fixations made on the foreground. The proportion of fixations on the foreground expected by chance factors alone are shown as a dashed lines with error bars indicating plus and minus one standard error of the mean.

It was found that the normal group fixated more frequently on the numerals when the task was to read them. Given that the images were displayed for 10 seconds each and that the numeral identification task is quite easy, it is possible that rather than rehearsing the numeral identity in short-term memory after the initial identification, participants continuously inspected the numeral as an active maintenance strategy. Consistent with this hypothesis, there was no significant increase in the proportion of fixations on the numeral for the colorblind group, as the colorblind participants had difficulty even identifying the numerals.

It is also reasonable to expect that the colorblind group would tend to search the images for visual cues that would aid in identification of the numeral, whereas normal controls would immediately identify the numeral and engage in detailed inspection behaviors. Consistent with this expectation, the colorblind group made more fixations, spent

less time fixating each location, and tended to make longer saccadic eye movements than the normal controls. These behaviors are indicative of visual search behavior.

2.4.1 Predicting Color Vision Status using Eye Movement Behavior

Given that there are observable differences in the eye movements of colorblind and normal individuals, these differences can be used as a metric for a color vision test. The differences are observable even when the participants are not explicitly instructed to identify the numerals as prescribed by the Ishihara test. This indicates that eye movements can be used as a covert metric to classify color vision status. To determine the effectiveness of this metric, the classification performance achievable was examined by considering the proportion of fixations made on the numerals. Of practical importance is that the classification performance was constructed without taking into account the other eye movement measures (e.g., fixation duration, saccade length) due to their wide range of observed values and relatively high variance in data. The wide range of observed values from these variables makes them less than ideal candidates for classifiers because they are prone to misclassification.

To calculate the classification performance curve on the basis of f*, the Bayesian decision theory was applied using the first *N* fixations observed in the experiment, where *N* ranged from 4 to 300 fixations. First, f* was calculated for each participant. Then, it was assumed that the population of f* would be normally distributed with mean μ and standard deviation σ in order that population distributions could be estimated from the observed f* of the normal group (n) and the colorblind group (c). The unbiased classification thresholds were then determined as the intersection between probability density functions of normal and colorblind group at each level of fixations (*N*) used in the analysis.

2.4.1.1 Optimum Statistical Classifier

This section explains how an unbiased classification threshold for normal and colorblind group was derived. From the Bayes rule (see [60]), the probability of an individual being in vision class w_i given that his proportion of fixation on the foreground, $\underline{x}$, has been measured can be defined as:

$$P(w_i \mid \underline{x}) = \frac{p(\underline{x} \mid w_i).P(w_i)}{p(\underline{x})} \tag{2.2}$$

where $\underline{x}$ is the continuous random variable that represents observed proportion of fixation on the foreground of an individuals, w_i represents vision class, and $P(w_i)$ represents the probability that class w_i will occur out of a population. The expression $p(\underline{x} \mid w_i)$ represents the probability density function for $\underline{x}$ given that we are measuring $\underline{x}$ from an individual whose vision class is w_i.

The Bayes classifier has decision function, $d_i(\underline{x})$, of the form

$$d_i(\underline{x}) = p(\underline{x} \mid w_i).P(w_i) \tag{2.3}$$

Because the probability density function $p(\underline{x} \mid w_i)$ are assumed to be Gaussian, the optimum decision function can be stated as

$$d_i(\underline{x}) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{\frac{(\underline{x}-\mu_i)^2}{2\sigma_i^2}}.P(w_i) \tag{2.4}$$

where σ_i is the population standard deviation and μ_i is the population mean of $\underline{x}$ from class w_i.

Accordingly, to calculated optimal classification threshold, or x, between normal (n) and colorblind (c) classes, the equation can be formulated as

$$d_n(\underline{x}) = d_c(\underline{x}) \tag{2.5}$$

or $$p(\underline{x} \mid w_n).P(w_n) = p(\underline{x} \mid w_c).P(w_c)$$

For simplicity of the analysis, the classification for this experiment was assumed that there is an equal likelihood that normal and colorblind classes occurring out of male population, yielding the equation

$$P(w_n) = P(w_c) \tag{2.6}$$

where $P(w_n)$ and $P(w_c)$ represents the probability of normal and colorblind individuals in male population respectively. Therefore, the unbiased decision boundary equation can defined as

$$p(\underline{x} \mid w_n).P(w_n) = p(\underline{x} \mid w_c).P(w_c)$$

or $$\frac{1}{s_n\sqrt{2\pi}} e^{-\frac{(x-m_n)^2}{2s_n^2}}.P(w_n) = \frac{1}{s_c\sqrt{2\pi}} e^{-\frac{(x-m_c)^2}{2s_c^2}}.P(w_c) \tag{2.7}$$

where x is the classification threshold, s_n and s_c are the standard deviation, and m_n and m_c are the mean of normal and colorblind group respectively.

By assuming that $P(w_n) = P(w_c)$, the optimal classification threshold, x, is the intersection of two probability density functions between normal and colorblind group can solved using the equation:

$$\frac{1}{s_n\sqrt{2\pi}} e^{-\frac{(x-m_n)^2}{2s_n^2}} = \frac{1}{s_c\sqrt{2\pi}} e^{-\frac{(x-m_c)^2}{2s_c^2}} \tag{2.8}$$

by applying a natural log, *ln*, for both sides of equation,

$$\ln\left(\frac{1}{s_n} e^{-\frac{(x-m_n)^2}{2s_n^2}}\right) = \ln\left(\frac{1}{s_c} e^{-\frac{(x-m_c)^2}{2s_c^2}}\right)$$

with ln*(a*b)* = ln *a* + ln *b*, and ln*(e)* = 1,

$$\ln(\frac{1}{s_n}) - \frac{(x-m_n)^2}{2{s_n}^2} = \ln(\frac{1}{s_c}) - \frac{(x-m_c)^2}{2{s_c}^2}$$

$$\ln 1 - \ln s_n - \frac{(x-m_n)^2}{2{s_n}^2} = \ln 1 - \ln s_c - \frac{(x-m_c)^2}{2{s_c}^2}$$

$$\frac{(x^2 - 2xm_n + {m_n}^2)}{2{s_n}^2} - \frac{(x^2 - 2xm_c + {m_c}^2)}{2{s_c}^2} + \ln s_n - \ln s_c = 0$$

by multiplying 2 on both sides of equation,

$$\frac{(x^2 - 2xm_n + {m_n}^2)}{{s_n}^2} - \frac{(x^2 - 2xm_c + {m_c}^2)}{{s_c}^2} + \ln {s_n}^2 - \ln {s_c}^2 = 0$$

$$(\frac{1}{{s_n}^2} - \frac{1}{{s_c}^2})x^2 - 2(\frac{m_n}{{s_n}^2} - \frac{m_c}{{s_c}^2})x + \frac{{m_n}^2}{{s_n}^2} - \frac{{m_c}^2}{{s_c}^2} + \ln {s_n}^2 - \ln {s_c}^2 = 0$$

From a quadratic equation $ax^2 + bx + c = 0$, x can be calculated from:

$$x = \frac{-b + \sqrt{(b^2 - 4ac)}}{2a} \tag{2.9}$$

by substituting

$$a = (\frac{1}{{s_n}^2} - \frac{1}{{s_c}^2})$$

$$b = -2 \times (\frac{m_n}{{s_n}^2} - \frac{m_c}{{s_c}^2})$$

$$c = \ln {s_n}^2 - \ln {s_c}^2 + \frac{{m_n}^2}{{s_n}^2} - \frac{{m_c}^2}{{s_c}^2}$$

where m_n and m_c are means, and s_n and s_c are standard deviations of the observed f* of the normal and colorblind group as they made *N* fixations. Figure 2.9 demonstrates optimal classification threshold, x, used to construct a classification performance that was quantified as the percent correct using the unbiased classification criteria. At each level of *N* fixations, an intersection of two distributions of the observed f* of the normal and colorblind

group, x_N, was used as optimal unbiased classification threshold. True Positive (TP or Hit) was the curve area of colorblind participants whose $f^* < x_N$, thus correctly classified as color defective. False Negative (FN or Miss) was the area of colorblind participants whose $f^* > x_N$. True Negative (TN or Correct rejection) was the area of normal participants whose $f^* > x_N$, thus correctly classified as normal. False Positive (FP or False alarm) was the area of normal participants whose $f^* < x_N$. Both FN and FP were areas of classification errors.

The percent correct at each level of *N* fixations was calculated from:

$$\textit{Percent Correct} = \frac{\textit{(True Positives + True Negatives).(100)}}{\textit{All population}} \qquad (2.10)$$

where *True Positives* were the number of colorblind participants whose $f^* < x_N$, and *True Negatives* were the number of normal participants whose $f^* > x_N$.

To construct the classification performance curve, first, at each level of fixations (*N*), the mean and standard deviation f* of each normal and colorblind individuals were calculated. These f* values represent corresponding distributions of f* of colorblind and normal group respectively at that level. The intersection, x_N, between the distribution of f* of normal and colorblind group was calculated using equation (2.9). Then true positives and true negatives were acquired by comparing the average value of f* for each individual with x_N at each level of fixations to calculate the percent correct by applying equation (2.10).

Figure 2.10 displays the classification performance curve that plots percent correct as a function of the total number of fixations used in the analysis. As can be seen in Figure 2.10, only 95 fixations are required to make a correct classification of color vision status in 95 percent of the cases. Given that the average number of fixations for each image was 33, 95 percent correct classification performance is achievable with the use of just three images.

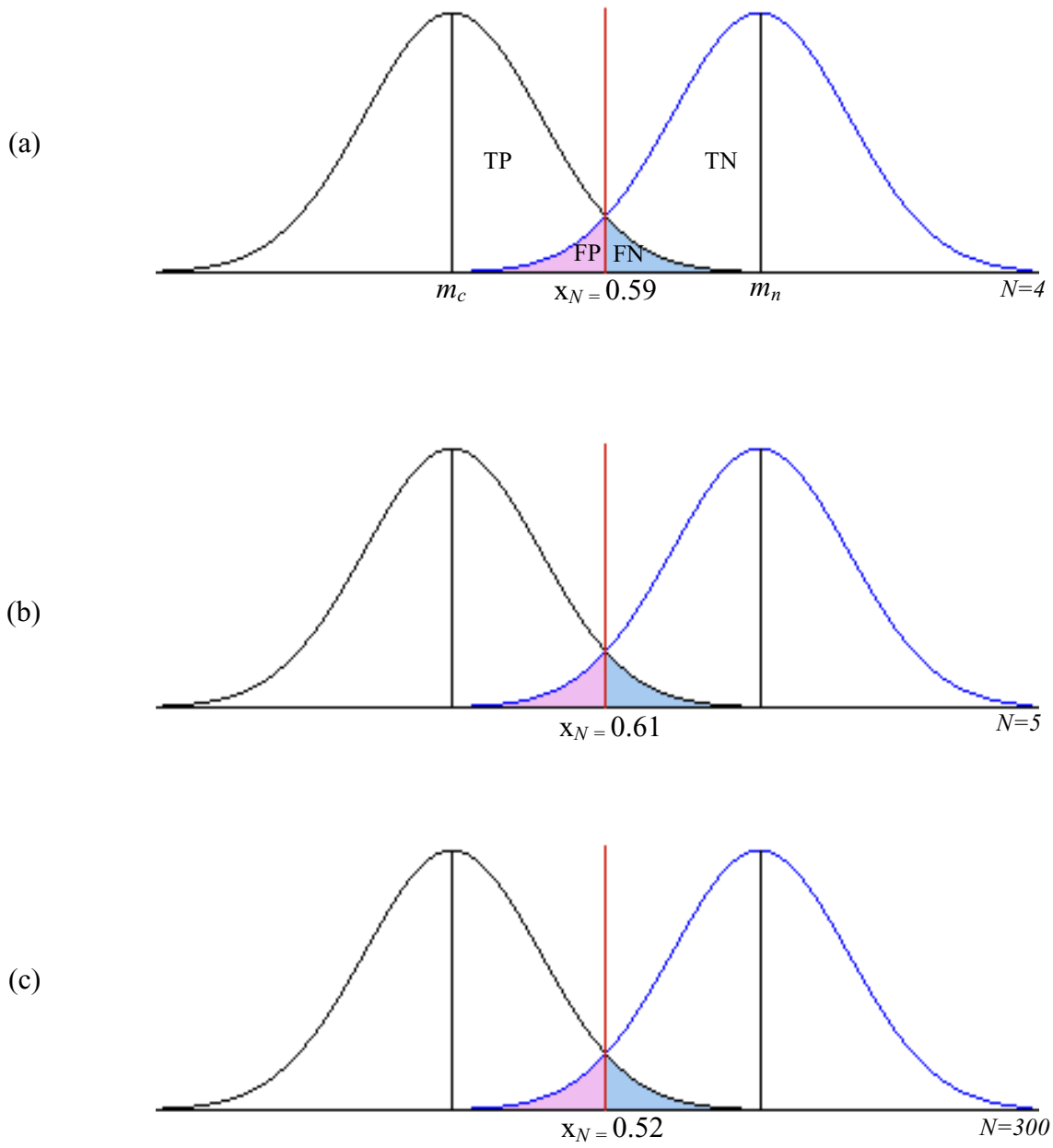


Figure 2.9. The normal distribution of f* of colorblind and normal group with mean m_c and m_n respectively as they made N fixations. N ranged from 4 to 300 fixations. At each level of N, the intersection, x_N, was used as an optimal unbiased classification threshold to calculate the percent correct. (a) $N = 4$ fixations; $x_N = 0.59$. (b) $N = 5$ fixations; $x_N = 0.61$. (c) $N = 300$ fixations; $x_N = 0.52$.

Note that each image was presented for ten seconds in the experiment. It is speculated that the efficiency of this metric could be further improved by presenting more images within the same period of time. Considering the time needed to administer this test, this approach requires less than one minute of eye tracking and the viewing of only three Ishihara images. The images do not necessary have to be viewed contiguously. This makes the test a very efficient method to detect color vision status.

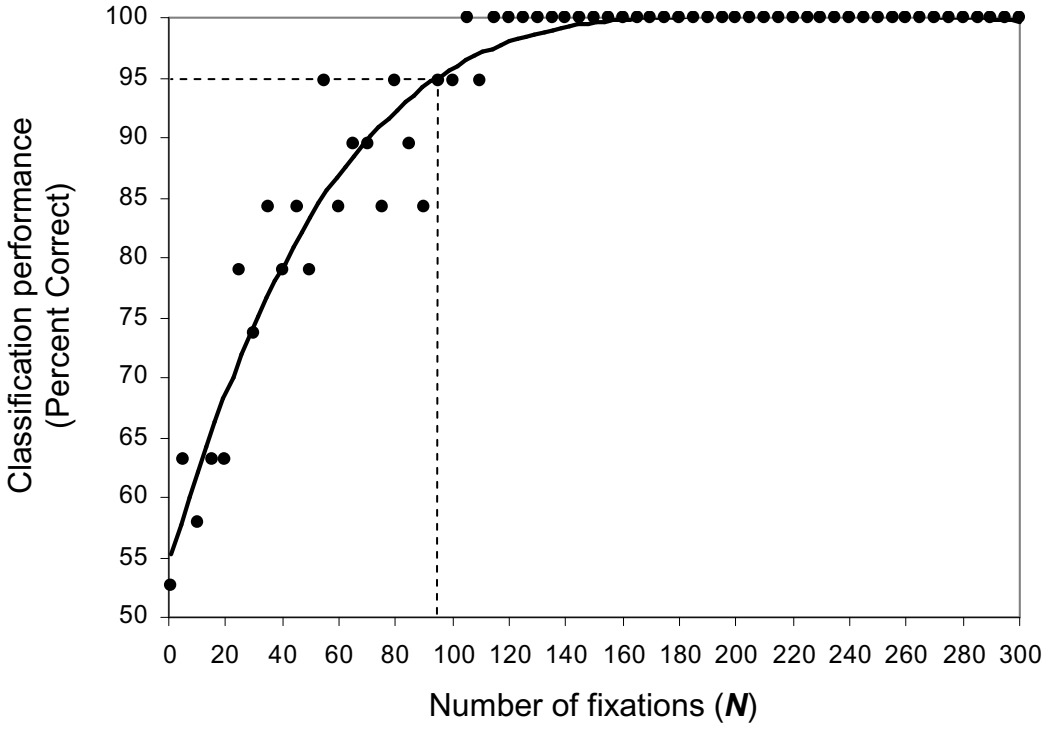


Figure 2.10. The classification performance curve indicates that only 95 fixations are required to obtain 95 percent correct color vision status classification accuracy.

Figure 2.11 shows the distribution of classification threshold (x_N) as a function of number of fixations taken from the experiment. As the number of fixations (N) increases, the amplitude of the classification threshold begins to decrease. The classification threshold appears approximately constant with respect to the number of fixations at around 190

fixations. The classification threshold eventually approaches the value of 0.52. On the basis of f* and *N* for the extended Ishihara test, this result implies that if an individual fixates on the foreground of the Ishihara images with average f* lower than 0.52, that individual will be classified as a colorblind. On the other hand, if such an average f* is higher than 0.52, that individual will be classified as a normal.

It should be noted that this result was based on the assumption that an equal probability of normal and colorblind vision status would occur in a population or $P(w_n) = P(w_c) = 0.50$ (see Equation 2.6). While this assumption was valid for the experiment, it is not for the population at large. Due to the fact that approximately 8.5% of the population (male and female combined) is colorblind, it would be practical to assume that 8.5% of the participants in any administration are colorblind when applying this test to a test population. However, a somewhat greater false positive (FP) rate and correspondingly smaller false negative (FN) rate for the screening test may occur due to this biased classification.

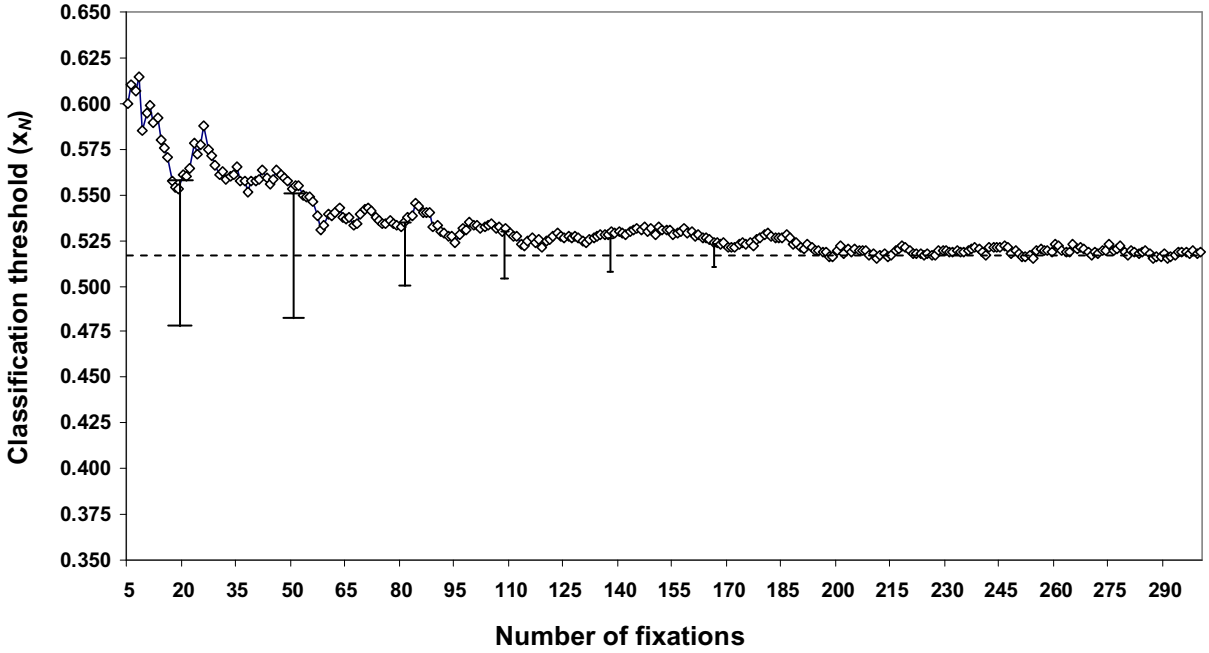


Figure 2.11. Relation between classification threshold (x_N) and number of fixations (N). As number of fixations increases, the classification threshold appears approximately constant, approaching the value of 0.52. The error bar represents varying amplitude of x_N at each interval between fixations.

2.4.2 Validation Test

To determine the qualitative diagnostic validity of this eye movement metric, color vision classification performance using f* was compared with that of the computerized Ishihara test [19]. The Cohen's Kappa (K) of agreement was used to establish the validity of the test. K was calculated as follows:

$$K = \frac{(P_O - P_C)}{(1 - P_C)} \tag{2.5}$$

where P_O is the relative observed proportion of agreement between two tests and P_C is the proportion of agreement predicted by chance.

This statistics has a value between –1 and 1 where K = 1 means perfect agreement, K = 0 means that agreement is due to chance, and K = –1 means perfect disagreement. Good

agreement between tests can be indicated if K is > 0.80 [94]. For an image of the Ishihara test, if an individual responded correctly, such a response was categorized as a normal (N) classification. Otherwise it was categorized as a colorblind (C) classification. Fixations drawn from the same image were used to calculate f*. If the f* was higher than the unbiased classification threshold (x_N), such an event was categorized as a normal (N) classification. Otherwise it was categorized as a colorblind (C) classification. It was found that the K coefficient of agreement between the response categories of the Ishihara test and the eye movement metric reached as high as 0.89 using as few as 5 images. Figure 2.12 shows the average K coefficient plotted as a function of the number of images used to calculate K.

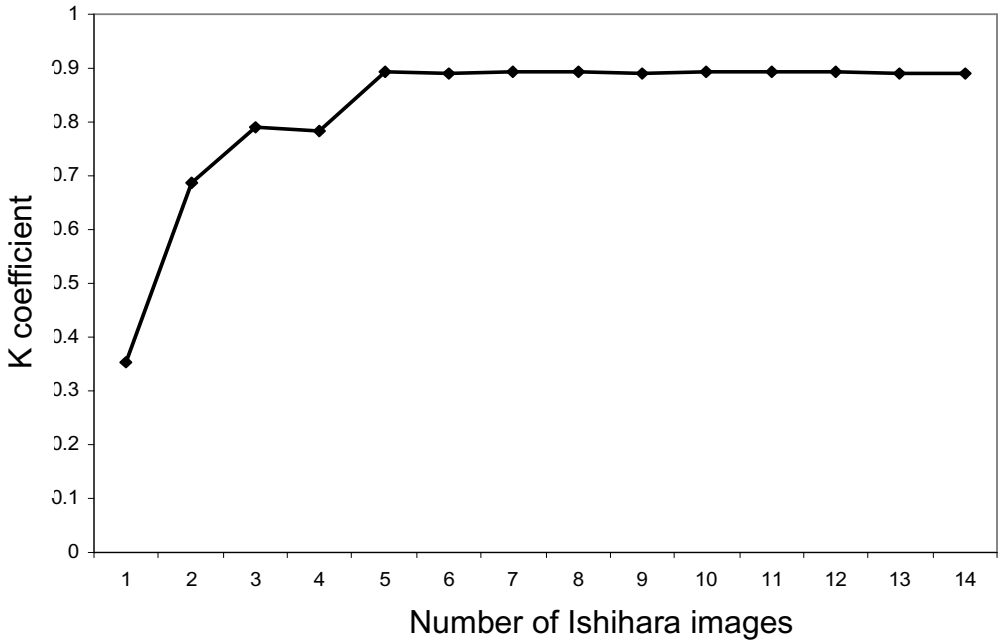


Figure 2.12. The average coefficient of agreement (K) was plotted as a function of the number of images used to calculate K. The agreement reached 0.8 using only 3 Ishihara images. Note that the level of agreement increased, approaching 0.89 when 5 images were used.

2.4.3 Sensitivity and Specificity of the Test

Another approach to validate the test is the analysis of sensitivity and specificity [61, 89]. Sensitivity is the ratio of true positives divided by the quantity of true positives and false negatives. Specificity is the ratio of true negatives divided by the quantity of true negatives and false positives [95]. A true positive occurs if a colorblind individual is classified as colorblind. A false negative is a case when a colorblind individual is classified as normal. Similarly, a true negative occurs if a normal individual is classified as normal whereas a false positive is a case when a normal individual is classified as colorblind (see Figure 2.9). For the Ishihara test, the sensitivity was calculated from the responses of the colorblind group, and specificity from the responses of the normal group. For example, it was counted as a true positive if a colorblind participant read an image incorrectly. Table 2.3 shows the sensitivity of the Ishihara test from 9 colorblind participants and the specificity from 10 normal participants from images 2-14.

For the eye movement metric, the sensitivity was calculated from the f* of the colorblind group, and specificity from the f* of the normal group. As shown in figure 2.9, at any N fixations, if the f* of a colorblind participant was lower than the unbiased classification threshold (x_N), such an event was categorized as a true positive. Otherwise it was categorized as a false negative. Table 2.3 demonstrates the sensitivity of the eye movement metric from nine colorblind participants and the specificity from ten normal participants as number of fixations increased. As the number of fixation reached 95 fixations, the sensitivity and the specificity of the eye movement metric was 0.88 and 1.0 respectively whereas percent of correct classification reached 95% (see Figure 2.10).

Table 2.3. Comparative results analysis of computerized Ishihara test vs. eye movement metric

Group	Computerized Ishihara test (14 images)			Eye movement metric (3 images or 95 fixations)		
	Response incorrectly	Response correctly	All responses	f*		Total participants
				Participants with f* < x	Participants with f* > x	
Colorblind	105 (TP)[a]	21 (FN)[b]	126	8 (TP)	1 (FN)	9
Sensitivity	0.83 (105/126)			0.88 (8/9)		
Normal	0 (FP)[c]	140 (TN)[d]	140	0 (FP)	10 (TN)	10
Specificity	1.00 (140/140)			1.00 (10/10)		
Percent correct classification[e]				95%		

[a] TP = True positive
[b] FN = False negative
[c] FP = False positive
[d] TN = True negative
[e] Percent correct classification = (TP+TN)/All participants = (10+8)/19

The screening efficiency of this test may vary depending on levels of defect severity. Similar to the Ishihara test, this test can more easily detect severe red-green color vision defects and does not attempt to classify the type of color-vision deficiency. A mild anomalous trichromat who experiences a minimal alteration of color perception might be able to pass this test.

In conclusion, this study revealed that there are differences in the eye movements of normally sighted individuals and individuals with color deficiencies. Given that stimulus properties are known to influence the guidance of eye movements, and that normal and colorblind individuals sense colors differently, it was predicted that these differences would be observed when participants viewed Ishihara test stimuli. These results and subsequent examination of the use of eye movements as a classification metric demonstrates the feasibility of extending the Ishihara test using eye movements.

CHAPTER 3. A METHOD FOR IMAGE AND REGIONS OF INTEREST SELECTION USEFUL FOR DIAGNOSING COLOR VISION DEFICIENCIES

The findings from the experiment in Chapter 2 suggested that as the Ishihara images were presented, colorblind individuals perceive numerals differently from those with normal color vision, and these differences can be detectable using an the eye movement metric. This metric also provides a novel approach to screen red-green color vision defects without relying on verbal communication. However, the Ishihara images are not the ideal candidate to be used as stimuli for covert screening application. As one of the most recognizable color vision tests, the Ishihara has been reprinted in numerous editions for over 90 years and has been used worldwide [24]. Using Ishihara images might allow participants to realize that their color vision is being tested. A problem exists when colorblind individuals who have seen the Ishihara images before may recognize the locations or correct shapes of the numerals [12]. The viewing strategy shifts from bottom-up to top-down control mechanisms, in which the participants' expectations influence visual attention.

A potential solution is to use photographs of complex scenes that are likely to be perceived differently by normal and colorblind observers as stimuli for a covert test. Complex scenes such as those with ordinary contents or everyday objects can be displayed without allowing the observers to realize the purpose of the test. The goal of this chapter is to describe a method to select photographs of complex scenes that are likely to be perceived differently by normal and colorblind observers. These images, when used as stimuli, should provide different visual influences between normal and colorblind observers. The logic is as follows.

Under natural viewing conditions, visual attention of normal observers is guided to regions with salient color contrast. Colorblind individuals have reduced sensitivity to red-green color contrast. Therefore, their attention should not be drawn to these regions. The focus of visual attention can be identified by direction of gaze [35, 80, 96]. Thus, observable differences in eye movements should be detectable for images with identifiable regions of red-green color contrast.

A method for selecting images with identifiable regions of red-green color contrast likely to lead to different patterns of gaze for normal and colorblind individuals is introduced. The regions of interest (ROIs) are regions where normal and colorblind individuals have different probability of fixating on each image.

Photographs of complex scenes have reportedly been used by optometrists and other professionals to better understand visual problems of colorblind patients. Spalding and Arden [97] published 22 color photographs of complex scenes in which red-green colorblind individuals could have problems perceiving the color objects. These images consist of both natural and architectural landscapes. Cold and Lain [13] later conducted an experiment using these photographs as stimuli and asking 79 colorblind and 20 color normal participants to identify color objects (i.e., berries and flowers) in natural settings. They found that all colorblind participants could locate fewer color objects than color normal participants. Unfortunately, the criteria on how to select these photographs were not publicly addressed [98]. Several studies have referred to the digital images with low red-green contrast that could cause colorblind individuals to misread the representation of visual information, but as well did not discuss how to determine which scenes would be perceived differently by people with normal and red-green colorblindness (i.e., see [99],[100]).

Some color images will be seen the same by everyone whereas some color images will be seen differently given that people with color vision deficiencies have diminished capacity to discriminate some colors in complex surroundings [10]. For example, colorblind individuals, regardless of type and severity of their defects, generally perform worse locating colored flowers and foliage in natural scenes than normally sighted people [13]. As shown on the left in Figure 3.1 are a pair of pictures that consist of blue sky and gray and brown structures. Individuals with red-green colorblindness are able to distinguish these colors. As a result, both normally sighted and colorblind individuals should share similar color experiences. On the other hand, the pictures on the right show green leaves and red berries. The red berries pop out and attract attention for normal individuals. Individuals with a moderate to severe red-green deficiency will perceive the red berries as brown because they are unable to distinguish between colors in the red-green section of the spectrum [101]. Thus the berries will blend into the background and fail to grab attention. The research described in this chapter was aimed at the development of a method to select digital images likely to be perceived differently by both groups, such as the one in this example.

Prior to the development of this method, there was no known way to rigorously determine which scenes would be perceived differently by people with normal and red-green colorblindness. The developed method establishes criteria to decide which images contain regions that could attract visual attention differently.

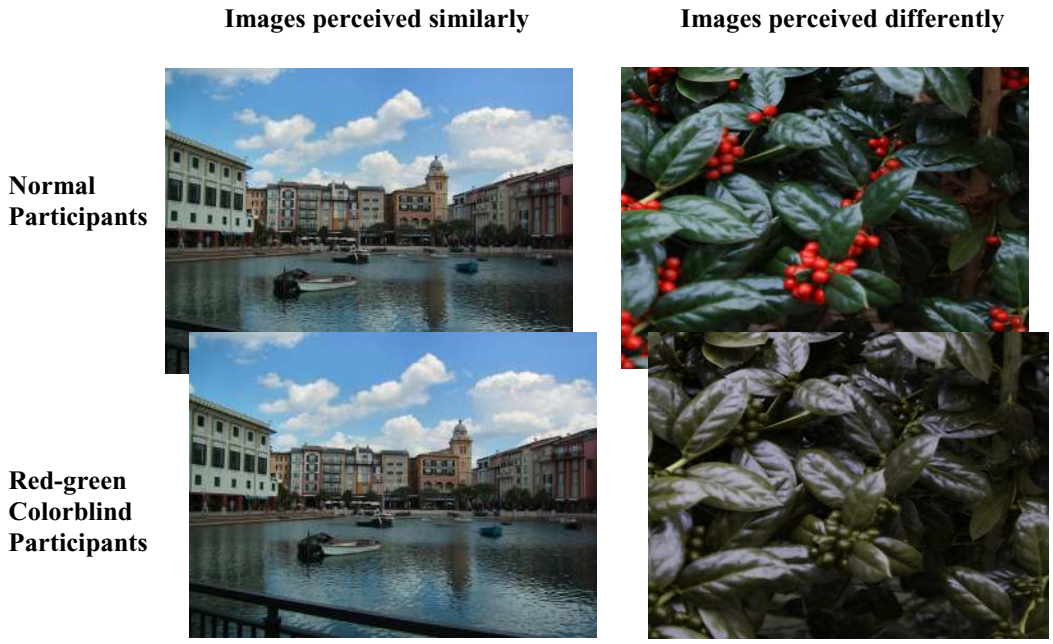


Figure 3.1. Simulation of color defective vision. On the left pairs of images, normal people should be able to perceive red hue on the berries while colorblind individuals would perceive red berries as dark brown berries. On the right pairs of images, both groups should share similar color experiences as colorblind individuals are not confused by blue, white, and greenish brown colors. The images in the second row were simulated as they would be seen by red-green colorblind observers using Meyer and Greenberg's dichromatic reproduction [88].

3.1 Overview of Method of Selecting Images and Identifying Regions of Interest

This method takes a random set of images as input and generates as output a subset of images with regions of interest. This method is designed to select photographs that are likely to be perceived differently by normal and colorblind observers. The method consists of six stages as diagramed in Figure 3.2.

Stage 1 takes a set of acquired images as input and generates a subset of color images that contain no people, text, or numerals as output. This filtering is implemented because images with people, text, or numerals are likely to attract visual attention in a way that is not purely stimulus driven. Stage 2 takes images acquired from Stage 1 as input and generates

colorblind version of input images as output using a dichromatic simulation technique. Stage 3 takes the normal images and colorblind images as input and generates four saliency maps (color saliency map, intensity saliency map, orientation saliency map, and final saliency map) for each image as output. The saliency model [75] is used to generate these saliency maps. This stage is implemented in order to predicted areas where normal and colorblind individual are likely to attend. Stage 4 takes the color saliency maps of the normal images and color saliency maps of its colorblind image as input and generates a color saliency difference map as output. This stage is processed by subtracting the color saliency maps of normal and colorblind image to obtain color saliency difference values. The positive salience values represent visible areas that are more attractive to normal than to colorblind individuals. The negative values represent visible areas more attractive to colorblind than to normal individuals. Stage 5 takes all normal images as input and generates a subset of images likely to be perceived differently by normal and colorblind individuals. In this stage, skew of color saliency difference is used as an index to rank the images with the most detectable areas of interest. Stage 6 takes color saliency difference maps of images from Stage 5 as input and generates corresponding +ROI maps and –ROI maps as output. The deviation from the mean of color saliency difference is used as thresholds to define regions of interest. If a value from color saliency difference map is higher (or lower) than threshold, its corresponding location of image is considered a +ROI (or –ROI). In this stage, the +ROI map displays regions of interest (ROIs) predicting areas where normal individuals are more likely to fixate than colorblind individuals. The –ROI map displays regions of interest predicting areas where colorblind individuals are more likely to fixate than normal individuals.

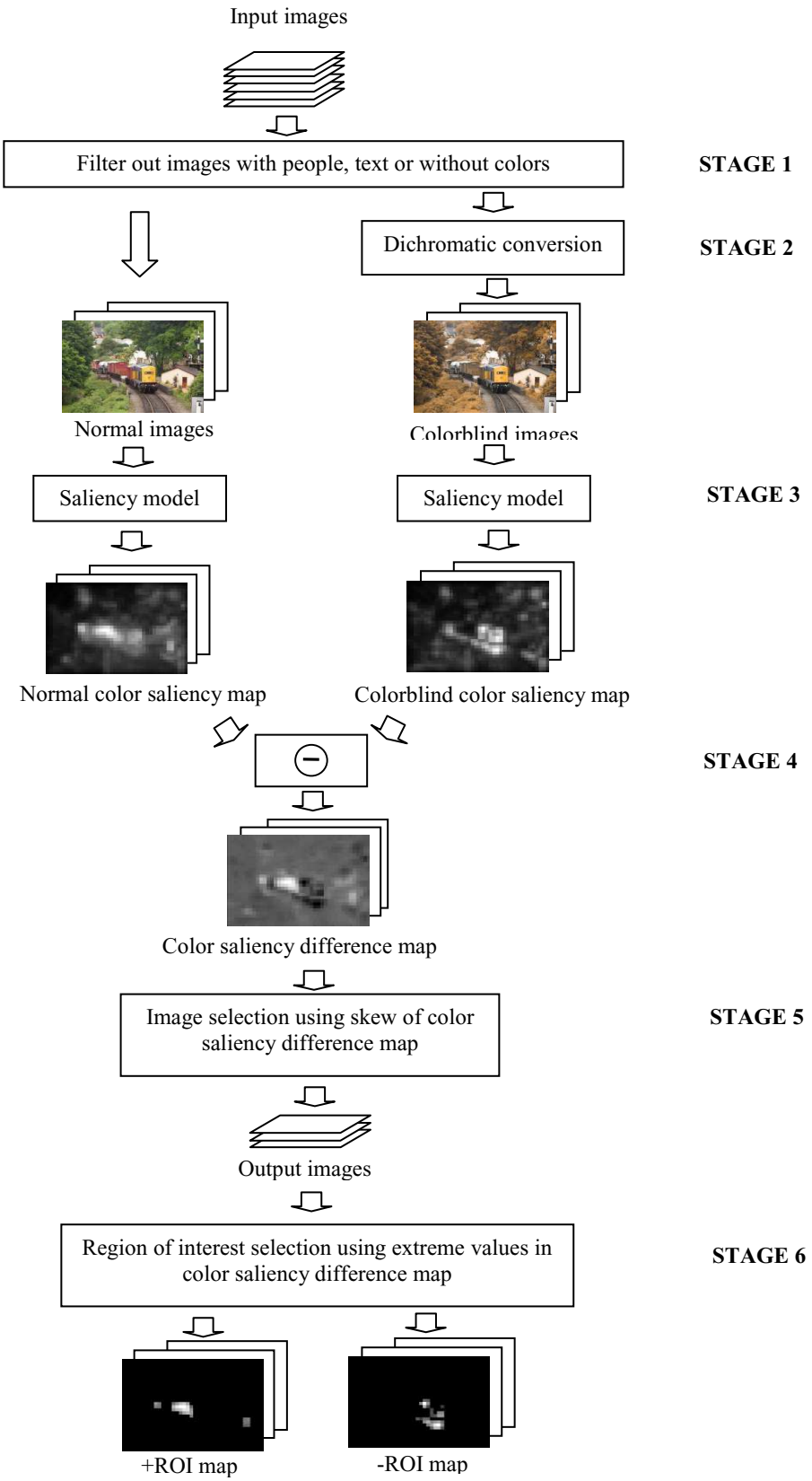


Figure 3.2. The method of selecting complex scene images and identifying ROIs.

3.2 Method

3.2.1 Stage 1: Manual Image Screening

This stage takes a set of acquired images as input and generates a subset of color images that contain no people, text, or numerals as output. In this stage, images that contained people, faces, skins, organs, text, or numerals are manually removed because these objects are likely to attract visual attention due to top down mechanisms, not driven by stimulus features [102, 103]. Monochrome images or images that contain only two colors, such as black-and-white, green-and-white or green-and-black are also removed so that digital images with a natural color depth can be used.

3.2.2 Stage 2: Simulation of Color Vision Defects

This stage takes color images acquired from Stage 1 as input and generates colorblind version of input images as output. Meyer and Greenberg's dichromatic algorithm [88] is used to simulate the colors as would be seen by red-green colorblind individuals. The input images from this stage are referred as "*normal images*" whereas the simulation images are referred to as "*colorblind images*" (See Figure 3.3).

(a) (b)

Figure 3.3. Example of (a) normal and (b) its colorblind image

To understand how color vision defect is simulated, the chromaticity diagram is useful. In the CIE XYZ diagram, there is a single point for each type of color deficiency, called *co-punctual point. Confusion lines* are straight lines radiating from co-punctual point [39]. Colors which are on confusion lines are those that people with red-green colorblindness get confused (see Figure 3.3). These colors along a confusion line look the same as all the other colors along that line.

Based on empirical studies, Meyer and Greenberg [88] assumed that color space of dichromats collapses to a line called "major axis" on the uniform chromaticity diagram (CIE LUV) for each of type of disorder (see Figure 3.4). For protanopes this axis connects 473nm and 574nm on the spectral locus with the white point (D65); for deuteranopes the line is very similar, connecting 477nm and 579nm. The point where the axis intersects the confusion line represents the perceived color for that confusion line (see Figure 3.5). As the perceived colors are inferable, the simulated image as would be perceived by red-green colorblind individuals can be constructed. Dichromatic deficiencies are considered a more severe form of defective color vision. Thus, color transformation designed to aid dichromats is also useful for deuteranomalous trichromats [101].

In this stage, the deuteranope model of colorblind simulation is used due to two reasons. First, those who have defective M photopigments (deuteranopes and deuteranomalous trichromats) constitute the largest group of colorblind individuals. Second, the protanopes suffer similar hue discrimination problems as the deuteranopes [1].

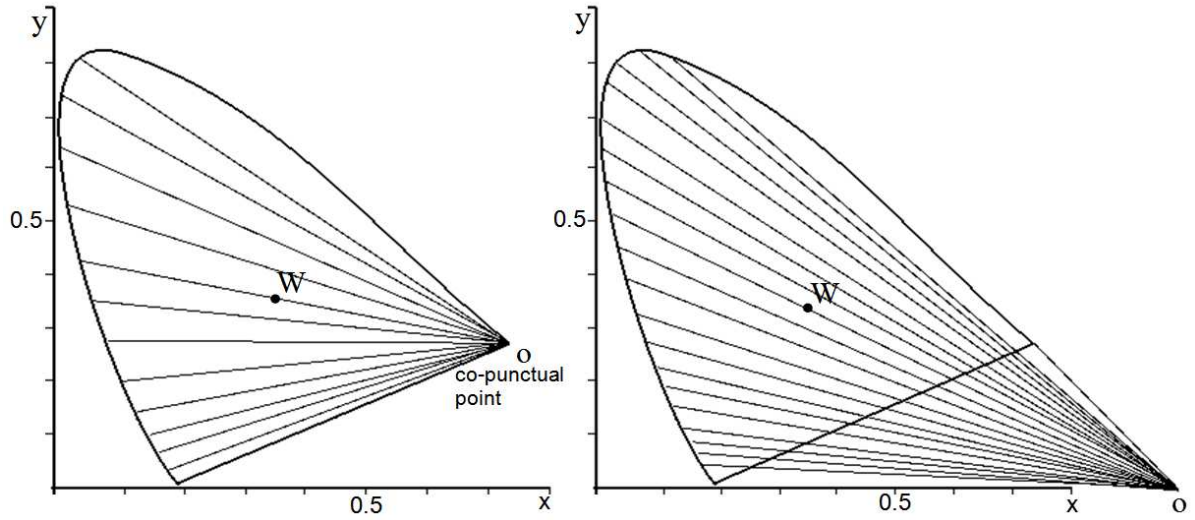


Figure 3.4. Confusion lines for (a) a protanope and (b) a deuteranope. Adapted from Schwartz (1999) [39].

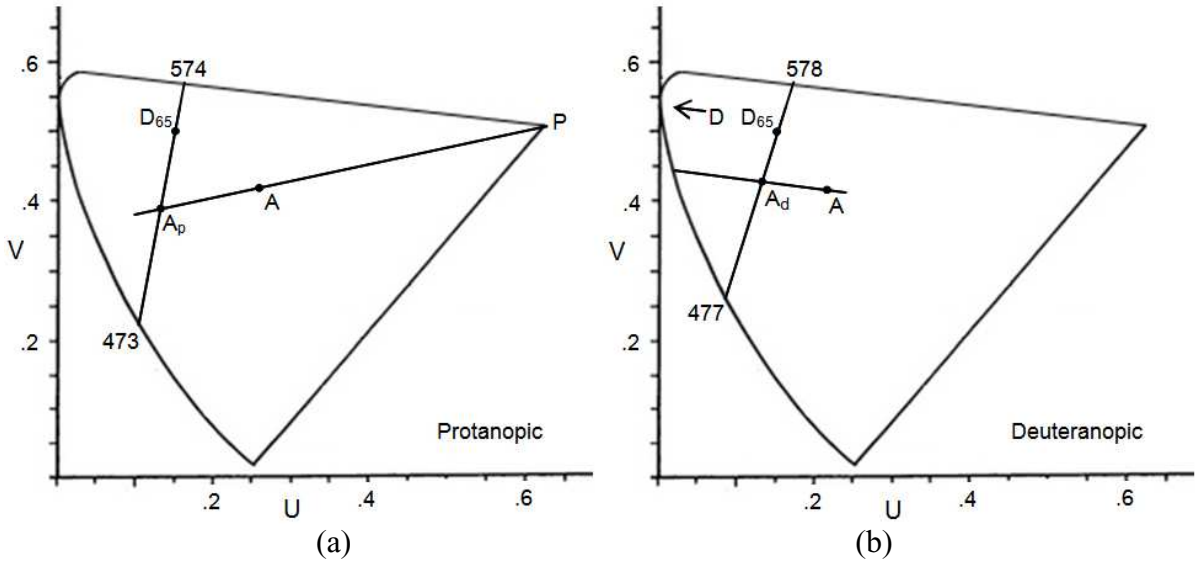


Figure 3.5. Axes of colors actually seen by dichromats on the CIE LUV color space. For example, the intersection between the confusion line (PA) and the major axis (473-574nm) represents color seen (Ap) by protanopes. Chromaticity points such as point A on the confusion line (PA) will be seen as chromaticity point Ap for protanopes. Adapted from Meyer and Greenberg (1988) [88].

3.2.3 Stage 3: Saliency Map

This stage takes each normal image acquired from Stage 3 as input and generates four saliency maps (color saliency map, intensity saliency map, orientation saliency map, and final saliency map) for each image as output. Then it does the same for each colorblind image. This stage takes an image as input and processes it through the saliency model [75] by decomposing it into 3 parallel features at different spatial scales. The output consists of four maps: three feature maps that encode areas which are found to be salient due to color, intensity, and orientation of the image and one final saliency map. The final saliency map is constructed by combining and normalizing the three feature maps. The map is plotted as a linear grayscale colormap with dimension of 40x30. This is an established scale similar to the resolution of visual attention in human observers [36]. The saliency map displays areas in the image that are most visually important. It is used to predict the focus of visual attention on an image. This stage is necessary because it provides the color saliency map predicting areas where normal and colorblind individuals are likely to fixate due to color feature of the image. The color saliency map of a normal image predicts where normal individuals tend to attend whereas that of a colorblind image predicts where colorblind individuals tend to attend.

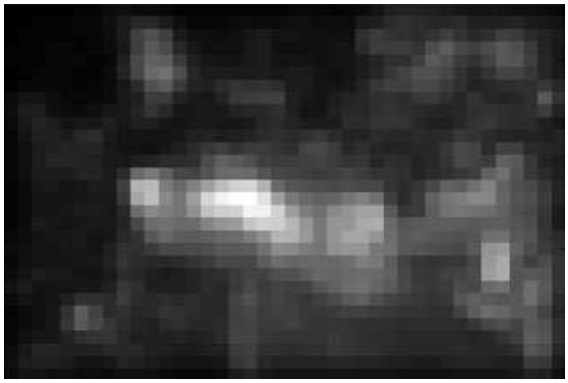
3.2.4 Stage 4: Color Saliency Difference Map

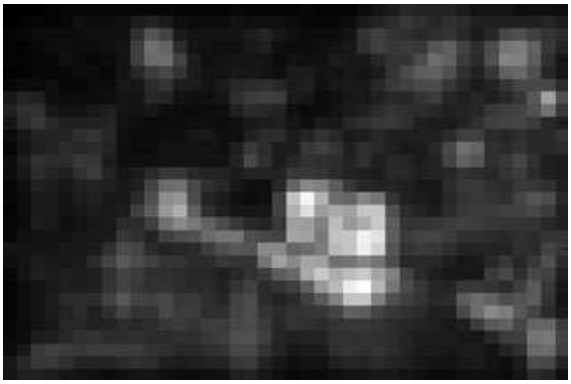
This stage takes color saliency maps of a normal image and color saliency maps of its colorblind image as input and generates a color saliency difference map for each image as output. In this stage, the color saliency map of a colorblind image is subtracted from that of a normal image to generate a color saliency difference map. In this stage, the color saliency map is used as input because it accounts for the color contrast in the image. The predicted areas of color saliency difference (ΔS) were calculated as follows,

$$\textit{Color saliency difference } (\Delta \boldsymbol{S}) = \boldsymbol{S_{nm}} - \boldsymbol{S_{cb}} \quad (3.1)$$

where $\boldsymbol{S_{nm}}$ is the matrix of color saliency values of a normal image, and $\boldsymbol{S_{cb}}$ is the matrix of color saliency values of a corresponding colorblind image. Brighter regions on the color saliency difference map represent regions that normal individuals are more likely to attend than colorblind individuals. On the other hand, darker regions represent regions that colorblind individuals are more likely to attend than normal individuals. The gray areas represent the regions that both groups could attend with equal opportunity.

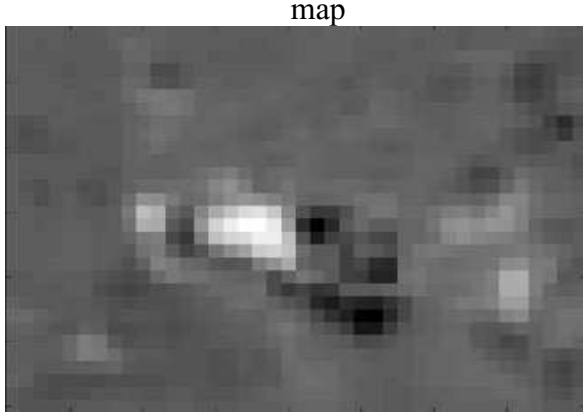
As demonstrate in Figure 3.6, the color saliency difference map of an image outlines the predicted salient areas. In this image, the engine of the train (dark region) is predicted as more salient to colorblind observers whereas the caboose (bright) is predicted as more salient to normally sighted observers.





(a) Normal image and its color saliency map

(b) Colorblind image and its color saliency map



(c) Color saliency difference map (ΔS)

Figure 3.6. Color saliency difference map between a normal and colorblind image pair displays predicted salient areas. Bright areas represent predicted areas salient to normal observers whereas dark areas represent predicted areas salient to colorblind observers.

3.1.5 Stage 5: Selecting images using skew of color saliency difference

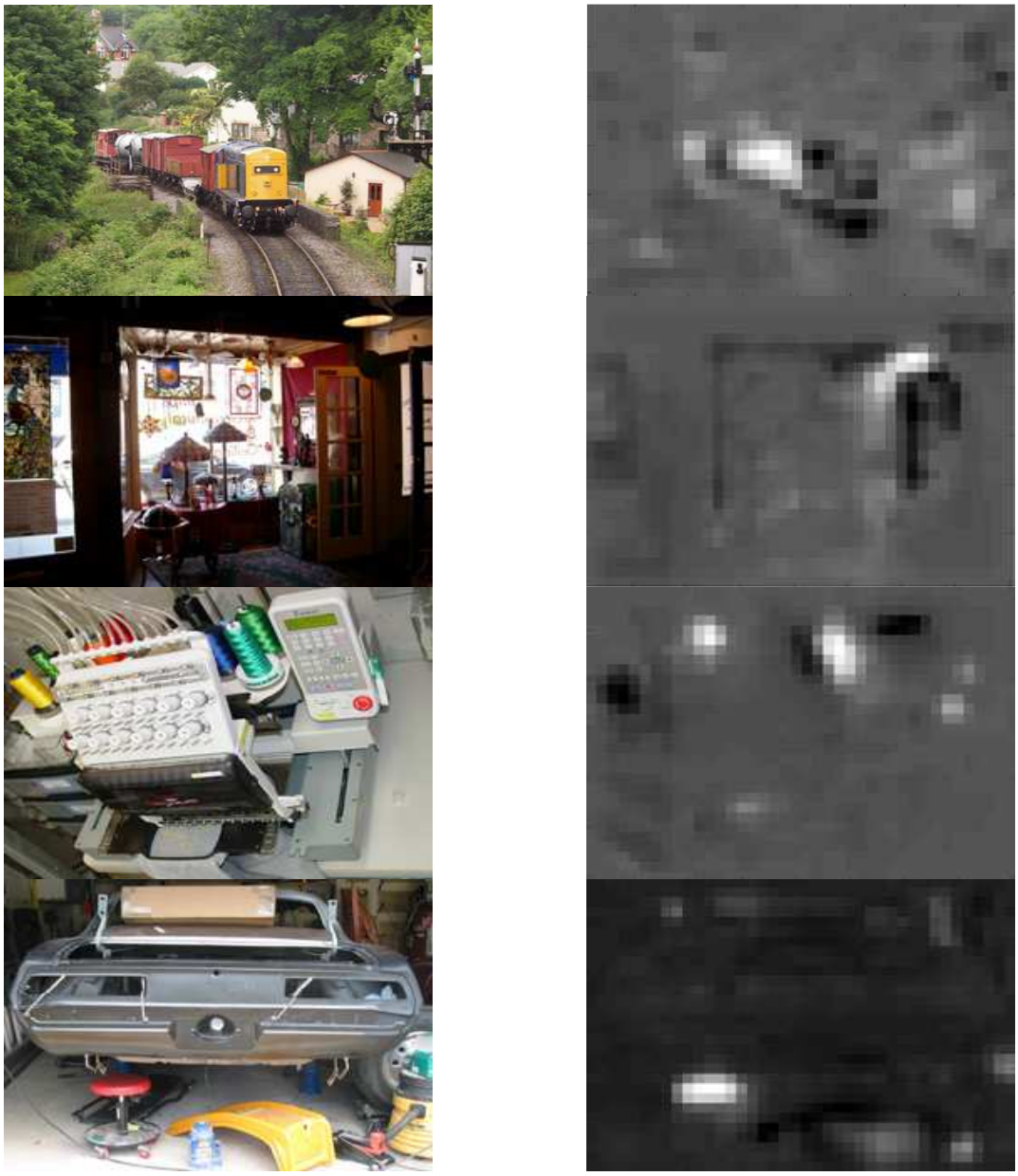
This stage takes all normal images as input and generates a subset of images with perceivable area of interest as output. In this stage, the skew of color saliency difference is used as an index to rank the images with visible clustered areas of interest. In an attempt to select images that best are represented as stimuli from all images, some color index is used for ranking images. More specifically, the index should help identify which image could provide clustered areas of interest that are perceptually as compact and as well separated as possible. The skew describes the deviation of the distribution of the saliency value from a Gaussian distribution. The absolute values of the difference between the normal and colorblind saliency map of each image is first calculated. The skew for each image is then obtained using the equation [59]:

$$skew = \frac{\sum_{x=1}^{W}\sum_{y=1}^{H}\left(\frac{S(x,y)-\overline{S}}{\sigma}\right)^3}{W \times H} \tag{3.5}$$

where $\overline{S}$ and σ representing the mean and the standard deviation of the saliency values *S(x,y)*, and *W* and *H* representing the width (*W=40*) and the height (*H=30*) of the map in pixels, respectively.

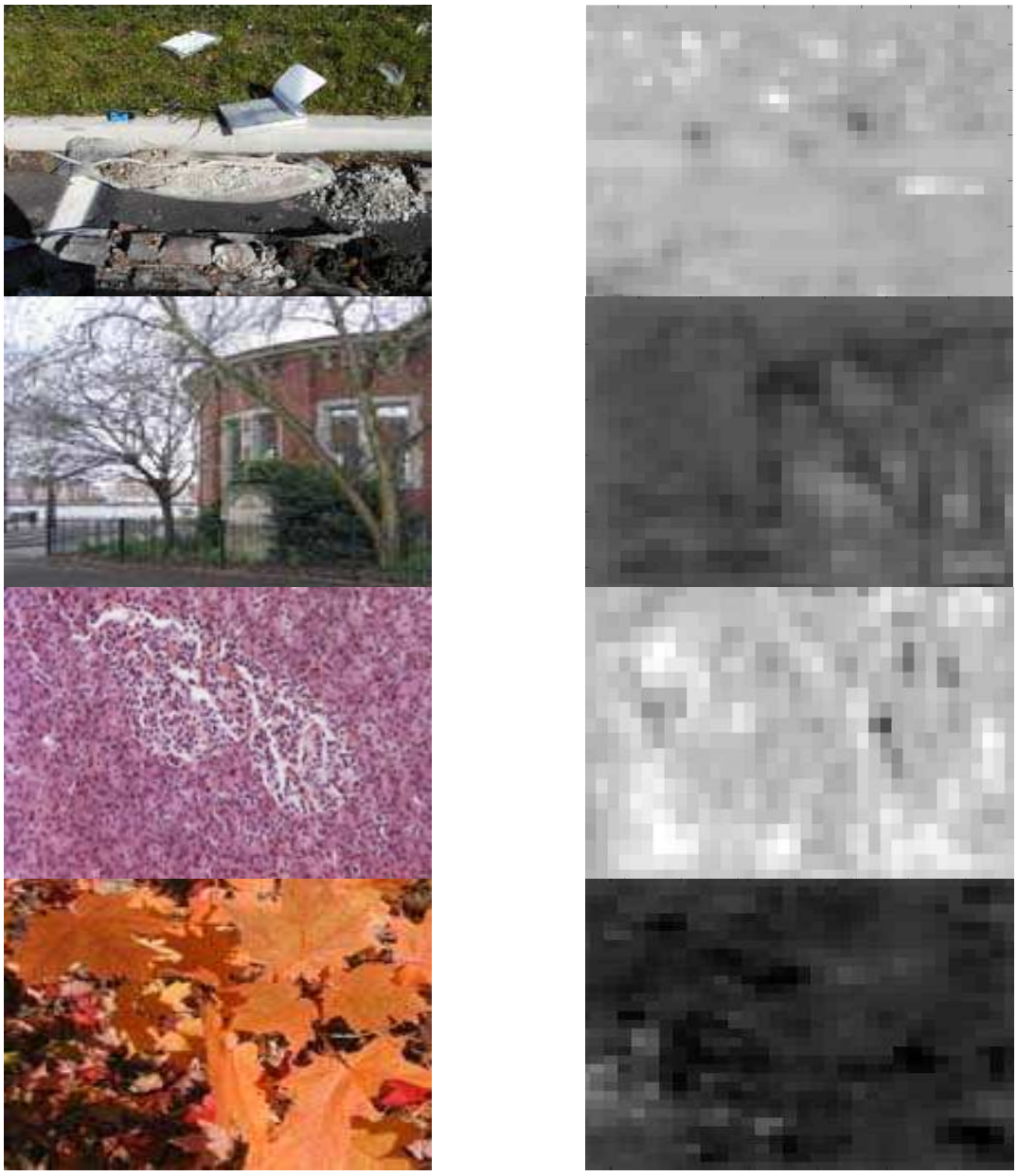
The higher value of the skew of color saliency difference, the better the clusters can be visibly identified from the surroundings. Demonstrated as an example in Figure 3.7, an image with higher skew of salience shows less scattering of bright area (e.g., clustered area obvious to normal observers) and dark areas (e.g., clustered area obvious to colorblind observers) than those with lower skew (see Figure 3.8). A possible explanation is that, in case of highly positive skew, the color saliency difference values were not normally distributed, demonstrating the asymmetry of those differences of an image. These high values rendered

visible clusters due to either positive difference (bright area) or negative difference (dark area) of salience values.



(a) (b)

Figure 3.7. (a) Example images with high skew of color saliency difference (value of 4.47, 4.52, 4.86, and 5.22 respectively). (b) Detectable areas of interest from color saliency difference map.



(a) (b)

Figure 3.8. (a) Example images with low skew of color saliency difference (value of 1.24, 1.40, 2.27, and 2.41 respectively). (b) Areas of interest from color saliency difference map are difficult to detect.

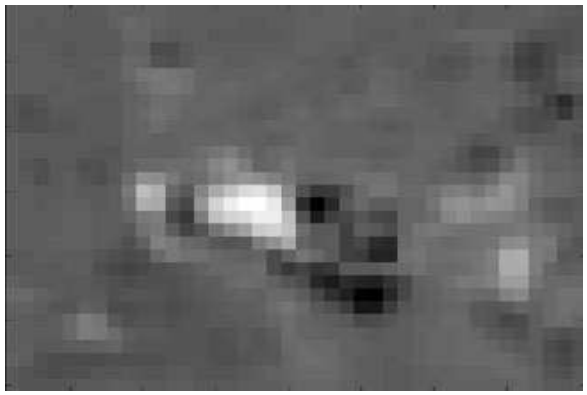
3.2.6 Stage 6: Identifying regions of interest

This stage takes color saliency difference maps of images selected from Stage 5 as input and generates regions of interest as output. In this stage, the +ROI map displays regions of interest to predict areas where normal individuals have higher chance to fixate more than colorblind individuals. The –ROI map displays regions of interest to predict areas where colorblind individuals have higher chance to fixate more than normal individuals.

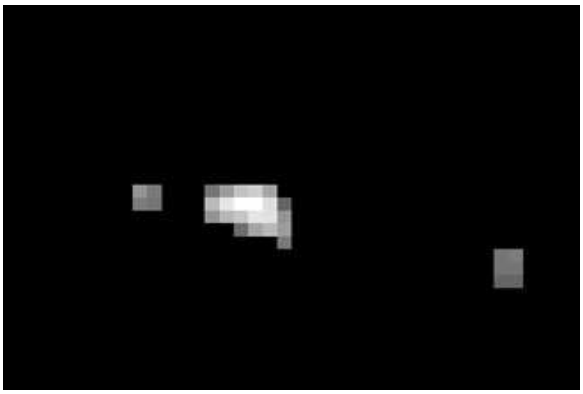
Regions of interest (ROI) are defined as areas with a large difference in the predicted eye movements between normal and colorblind individuals. The deviation from the mean color saliency difference (ΔS) is used to define upper and lower thresholds. It is assumed that the regions defined by extreme values would best distinguish between normal and colorblind individuals. The predicted ROIs are calculated from the input color map pair as follows,

$$\begin{aligned} S_{(i,j)} &= +ROI \quad if\ S_{(i,j)} > +\delta SD(\Delta S) \\ S_{(i,j)} &= -ROI \quad if\ S_{(i,j)} < -\delta SD(\Delta S) \end{aligned} \tag{3.2}$$

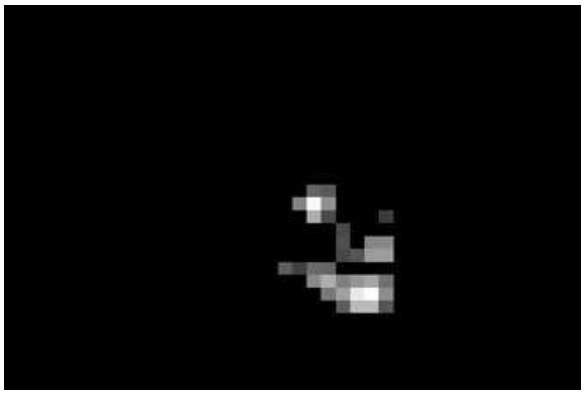
where $S_{(i,j)}$ is the saliency value of pixel (i,j) from the difference of color map between the normal and colorblind image (ΔS), and an δ is a threshold of deviation. The +ROIs are the pixels that contained saliency values higher than δ standard deviations, and the –ROIs are the pixels with saliency values less than negative δ standard deviations from the areas of saliency difference (ΔS). Figure 3.9 demonstrates the ROIs determined from a threshold value of δ = 3.



(a)

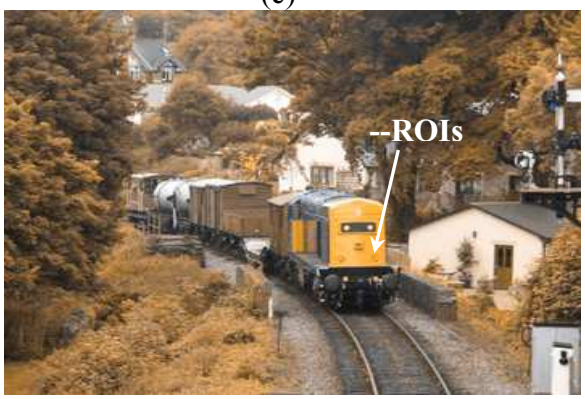


(b)



(c)

(d)



(e)

Figure 3.9. Example of image whose ROI defined as ± δ SD from Δ*S* (δ = 3). (a) Color saliency difference map (Δ*S*). (b) Regions of interest for normal individuals (+ROIs). (c) Regions of interest for colorblind individuals (–ROIs). (d) +ROIs (i.e., red caboose and door). (e) –ROIs (i.e., yellow train engine).

Since predicted ROIs are derived from the difference of color channel from a normal and colorblind image, such a difference has influence on the color of the ROIs. As previously discussed, the Meyer and Greenberg's dichromatic conversion computes color seen by colorblind individuals by replacing confusing colors (e.g.. shades of red and green) with the intersection between the confusion line and the major axis (e.g., shades of yellow). Therefore, red and green shades are basically removed out of the colorblind image. The color saliency map of the colorblind image would contain no saliency values due to contrast of red or green. The difference between the normal and colorblind map (ΔS) could yield positive values representing saliency due to color contrast most obvious to normal individuals, zero value representing saliency due to color contrast obvious to both groups, and negative values representing saliency due to color contrast most obvious to colorblind individuals.

3.3 Results

From an unbiased sample of approximately 23,000 digital images randomly and automatically downloaded from the Internet, 100 images were selected by using the described method (Appendix B). The regions of interest were defined with $\delta = 3$ (3SD). As the images were selected, the highest skew value was 6.48 while the lowest skew value was 3.96. The average size of +ROIs was 1.60% of the total image area with the standard deviation of 0.78. The average size of the –ROIs was 0.87% of total image area with the standard deviation of 0.65. Examples of the selected images are demonstrated in Figure 3.10 with colorblind image and the +ROIs and –ROIs plotted on the same map.

(a) (b) (c)

Figure 3.10. Example of selected images with high skew values. (a) normal image (b) colorblind image (c) +ROIs (bright regions or regions that normal individuals are likely to attend) and –ROIs (dark regions or regions that colorblind individuals are likely to attend).

3.4 Summary

This chapter outlined a method of selecting digital photographs that predict different patterns of visual attention for normal and colorblind observers. In addition, a method to identify specific regions of interest which compete for visual attention from normal and colorblind individuals was described. This process can be summarized in 6 stages. First, images with humans, text, and numerals are filtered out. Second, the remaining images are processed with color dichromatic conversion. This stage generates *colorblind images* from the original input. Third, the difference of color contrast between the normal and the colorblind image is taken into account by the saliency model whose color feature map is used to present influence of saliency due to color contrast. Fourth, the skew index is a suggested index used to rank the images with visible clustered areas. Finally, a guideline to define the region of interest for normal and colorblind individuals is described. The +ROIs are regions where normal individuals have higher chance to fixate more often than colorblind individuals. The –ROI map displays regions where colorblind individuals have higher chance to fixate more often than normal individuals.

More than 23,000 digital images were randomly downloaded from the Internet and 100 images were selected by using the described method. From the selected images, the highest skew value was 6.48 while the lowest skew value was 3.96. The average size of +ROIs was 1.60% of the total image area while that of the –ROIs was 0.87% of total image area.

CHAPTER 4. EXPERIMENTAL VALIDATION OF THE IMAGE AND REGION OF INTEREST SELECTION METHOD

The goal of the research presented in this chapter is to validate the method to select photographs of complex scenes that are likely to be perceived differently by normal and colorblind individuals. It was revealed in Chapter 2 that there are differences in the eye movements of normal and colorblind individuals when viewing the Ishihara images. The findings suggested that the Ishihara images can influence the allocation of attention of normal and colorblind individuals differently. Given that the selected photographs can influence the allocation of attention as well, it was predicted that the differences of visual attention would be observed when normal and colorblind individuals view these photographs. In this experiment, eye movements were used to quantitatively assess visual attention for the selected images. The predicted regions of interest (ROI) on each image were defined as target areas that attract visual attention from participants. The logical assumptions underlining the experiment are addressed as follows.

Under natural viewing conditions, visual attention of the observers is guided to regions with highest saliency. Normal and colorblind individuals are visually attracted to different regions of interest because their ability to detect color contrast is different. The focus of visual attention can be identified by direction of gaze [35, 80, 96]. Thus, observable differences in eye movements as measured by fixations between groups should be detectable in both the +ROIs and –ROIs. The +ROIs are predicted to be fixated more by the normal individuals whereas the –ROIs are predicted to be fixated more by the colorblind individuals.

4.2 Methods

4.2.1 Stimulus

To obtain an unbiased sample of the images, approximately 23,000 digital images were randomly and automatically downloaded from Yahoo.com basing on keyword search. The programming script written with Perl was executed on a computer server for 48 hours where it randomly selected keywords from a dictionary and obtain large size images related to those keywords from the web site. The acquired images were in the JPEG format with the smallest resolution of 640x480.

Images that contained people, faces, skin, organs, texts, and numbers were removed with a MATLAB algorithm since these objects are more likely to attract visual attention from humans regardless of the images' surroundings [102, 103]. Monochrome images were also screened out during this process so that only color images were used. These processes reduced the number of images from approximately 23,000 to 5,639 images. These images were manually checked to ensure that they provided adequate detail when displayed full screen on the monitor.

Then, the remaining images were processed with the method for selecting images and ROIs (see Chapter 3). First, they were processed with colorblind simulation algorithm to generate normal and colorblind images. Then each image was processed with the saliency model. Third, the color saliency maps of a normal and its corresponding image were subtracted to obtain the color saliency difference map. Fourth, total of 100 images with the highest skew of the color saliency difference were selected as stimuli in this experiment. Finally, the ROIs were defined for target areas for each image. In this experiment, the $\delta = 3$ (3SD) was selected as deviation factor from the color saliency difference map. This was

based on an empirical rule that extreme values generally are values higher than 3 standard deviations in the normal distribution.

4.2.2 Participants

Twenty male students from Iowa State University participated in the experiment. A statement of informed human consent was obtained from all the participants. Each participant was paid $10 for participation. Ten participants were selected on the basis of having reported that they were clinically diagnosed as red-green colorblind in the past. The remaining ten participants were selected as a comparison control group based on a self-report of normal color vision. To confirm their color vision status, all participants were administered with the computerized Ishihara test after the eye movement experiment. The interpretations for participants' color vision status from the Ishihara test results were all in accordance with the self-report of the participants.

4.2.3 Apparatus

Similar to the first study, all selected images were digitally adjusted to display full screen at a resolution of 1024 x 768 pixels in the 32-bit color mode and presented on a 19-inch monitor. The center of the display was aligned at the participant's eye level at a distance of 75 centimeters. This provided an optical image subtending 27˚ horizontally and 24˚ vertically that facilitate the tracking of eye movements. The established procedure for utilizing the Ishihara color vision test on a computer screen as well as the monitor calibration for the experiment were the same as the first study.

4.2.4 Procedures

The experiment contained two five minute blocks separated by a three minute break. Each block presented a series of 50 selected images. For both blocks, participants were seated in front of the computer screen with their head resting on a chin rest. At the beginning and the end of each block of the experiment, participants were required to fixate nine points on a full-screen grid. These measurements were required in order to calibrate the eye tracker and determine the average accuracy of the eye movement measurements. The average eye tracking error at the beginning of the block was 1.15˚ while average error at the end was 1.21˚. A t-test comparison indicated that there was no significant difference in average error of eye tracking between groups ($t(18) = 1.07$, $p = 0.291$).

During the experiment, each image was presented for a period of five seconds. Between each image presentation, a black dot at the center of the screen was presented for one second to be fixated prior to each trial. For both blocks of the experiment, participants were instructed to freely look around at the images. After the eye tracking process was completed, each participant was administered with the computerized Ishihara color vision test [89] to confirm color vision status. The results of this test confirmed the self-reported color vision status in all cases.

4.3 Data Analysis

ROI (–ROI, +ROI) and vision (normal, colorblind) were independent variables in this study. To quantify the number of fixations a participant made in the +ROIs, the percentage (f^*_{+ROI}) of the number of fixations in every +ROI (f_{+ROI}) relative to those on the

total image area ($\sum f$) was calculated for each image as $f^*_{+ROI} = f_{+ROI} / \sum f$. Similarly, the percentage of the number of fixations in every –ROI was defined as $f^*_{-ROI} = f_{-ROI} / \sum f$.

To estimate the percentage of fixations on the +ROIs expected by chance alone ($f_c^*_{+ROI}$), the percentage of areas of all +ROIs to the total area was calculated. The same notion was applied to the percentage of fixations on the –ROIs expected by chance alone ($f_c^*_{-ROI}$). The $f_c^*_{+ROI}$ across all images was 1.60% whereas the $f_c^*_{-ROI}$ across all images was 0.87%.

Since participants tend to fixate in central regions of the images by nature, an image shuffling technique was used to estimate chance performance that accounts for this central bias [36, 37]. This technique randomly shuffles participants' observed fixation pattern onto different images in order to disassociate fixation locations and image content while retaining the distributional properties of the fixations (i.e., centrally biased). If this more conservative chance performance baseline and the f^*_{ROI} differ, it suggests that participants' fixations are indeed guided by image properties, and are not due solely to the greater chance of central fixation.

Accordingly, the conservative estimate of the f^*_{ROI} expected by chance alone ($f_s^*_{ROI}$) can be calculated by overlaying each participant's observed fixation locations of an image onto the other images. This was achieved by a MATLAB script which overlaid each fixation pattern on each of the other 99 images, and calculated the average of $f_s^*_{ROI}$. For example, the fixation pattern of the first image was overlaid on the second image, third image, and so on. Then the fixation pattern of the second images was overlaid on the first image, third image, and so on. This shuffling process continued until the fixation pattern of the one hundredth

images was overlaid on the first image, second image, and so on. For each shuffle, a new value of f_s*_{ROI} for both +ROIs and –ROIs was calculated.

With this technique, the average f_s* in +ROIs across all participants (f_s*_{+ROI}) was 1.98% while the average f_s* in –ROIs across all participants (f_s*_{-ROI}) was 1.10%. These values are higher and more conservative than f_c*_{ROI}, therefore, they were used to estimate the chance performance in this study. Four eye movement measures: the percentage of fixations on the ROIs ($f*_{ROI}$), the average fixation duration, the average number of fixations per image, and the average saccade length were analyzed in this study. Each dependent measure was analyzed using a repeated-measures ANOVA with ROI (+ROI vs. –ROI) as a within-participant factor and vision (normal vs. colorblind) as a between-participant factor.

4.4 Results

Percentage of fixation on ROIs ($f*_{ROI}$) The average $f*_{+ROI}$ of the normal group was 5.95% (SD = 1.90%) whereas for the colorblind group it was 2.45% (SD = 0.65%). On the other hand, the average $f*_{-ROI}$ of the normal group was 1.18% (SD = 0.38%) whereas for the colorblind group it was 1.95% (SD= 0.83%). This is shown in Figure 4.1. Both main effects of vision and the ROI were found significant. A significant interaction effect ($F(1, 18) = 45.80$, $p < .05$) revealed that effect of ROI depended upon the vision status. A post-hoc test using a Bonferroni adjustment ($\alpha = 0.05$) revealed that when the images were presented, the normal group made more fixations on the +ROIs than the colorblind group did. In contrast, the colorblind group made more fixations on the –ROIs than the normal group did.

A t-test comparison indicated that the average $f*_{ROI}$ on the –ROIs of colorblind groups ($f*_{-ROI} = 1.95$) were significantly higher than that expected by chance factor alone

(f_s*_{-ROI} = 1.10), ($t(9) = 3.21$, $p < .05$). This was not the case for the normal group ($f*_{-ROI}$ = 1.18) as their percentage of fixations on the –ROIs did not differ from the chance level. However, the average $f*_{ROI}$ on the +ROIs of both the normal group ($f*_{+ROI}$ = 5.95) and the colorblind group ($f*_{+ROI}$ = 2.45) were significantly higher than that expected by chance factors alone (f_s*_{+ROI} = 1.98).

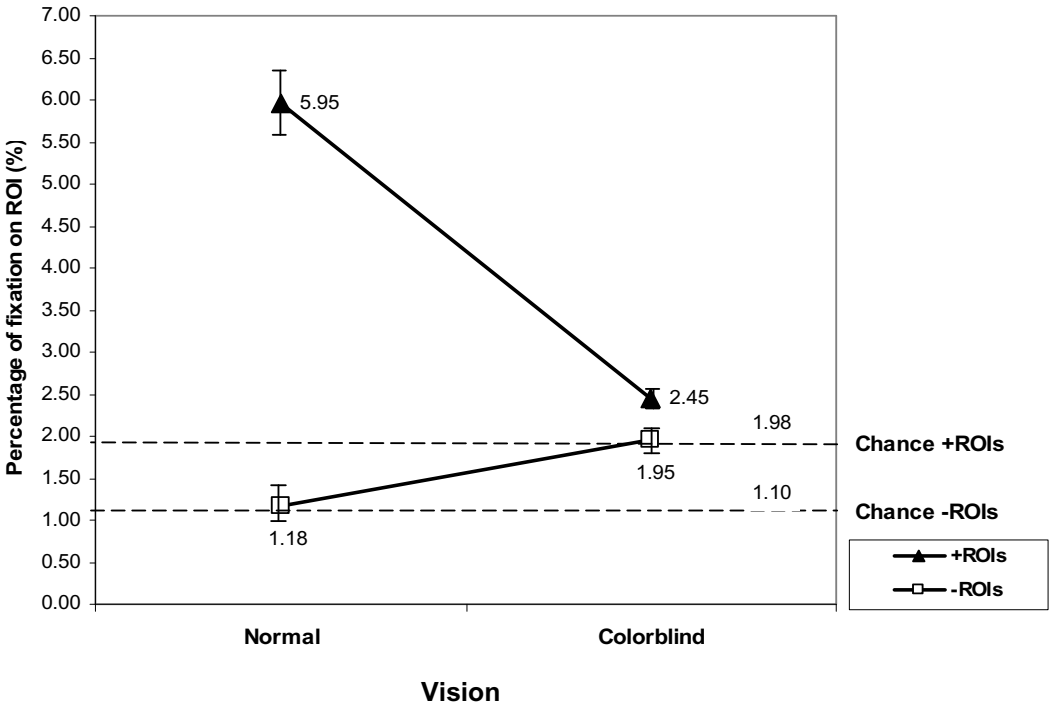


Figure 4.1. The percentage of fixations in the +ROIs was significantly lower for colorblind participants than normal controls. However, their percentage of fixations in the –ROIs was significantly higher than normal controls. The percentage of fixations on the foreground expected by chance factors alone (f_s*_{ROI}) are shown as a dashed lines.

Fixation Duration The average fixation duration of the normal group was 0.31 seconds (SD = 0.20 seconds) on +ROIs and 0.24 seconds (SD = 0.23 seconds) on –ROIs whereas for the colorblind group it was 0.27 seconds (SD = 0.22 seconds) on +ROIs and 0.29 seconds (SD = 0.19 seconds) on –ROIs. There was no significant main effect for either vision

or ROI. There was a significant interaction between vision and ROI ($F(1, 18) = 12.495$, $p = 0.002$) as shown in Figure 4.2. A post-hoc test using a Bonferroni adjustment suggested that the normal group spent more time fixating on the +ROIs but less time when fixating on the –ROIs as well as on the remaining areas ($p<0.05$). However, there was no significant difference among the amount of time the colorblind group spent on +ROIs , –ROIs, or the remaining areas.

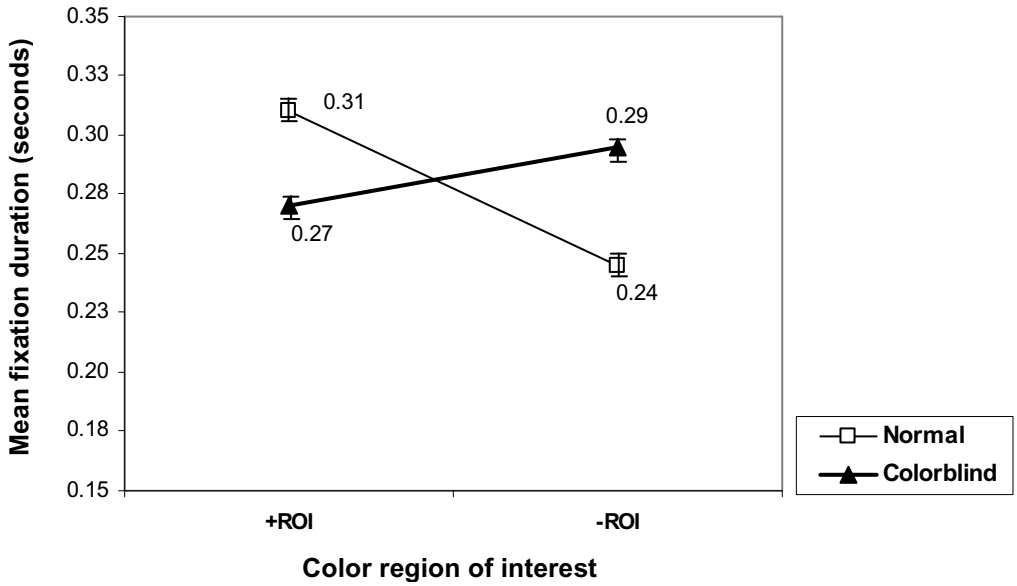


Figure 4.2. Average fixation duration in the ROIs. The significant interaction between vision and ROIs indicates that ROIs had an effect on participants, but its effect depended on their visual conditions.

Number of Fixations The average number of fixations per image of the normal group was 14.75 (SD = 1.75) whereas for the colorblind group it was 14.29 (SD = 1.55). No significant difference in the average number of fixations per image between groups was found. The average number of fixations of the normal group was 8.7 on +ROIs, 1.8 on –ROIs, and 13.7 on the rest whereas for the colorblind group it was only 3.5 on +ROIs, 2.8 on –ROIs, and 13.6 on the rest. It was found that normal participants landed more fixations on

+ROIs than colorblind participants on almost every image, with the exception of image no.1 and 23 where both groups made the same total number of fixations, and image nos.11,35,41,43,66,67,70,77, and 93 where both groups made no fixation on +ROIs. Consistent with the ANOVA results of the percentage of fixation, a post-hoc test using a Bonferroni adjustment revealed that the normally sighted group fixated more on the +ROIs than did the colorblind group ($p<0.05$).

Saccade Length The average saccade length of the normal group was 5.13 degrees (SD = 1.62 degrees) whereas for the colorblind group it was 4.98 degrees (SD = 1.24 degrees). No significant difference in the average saccade length between groups was found.

4.5 Discussion

The purpose of this study was to test the hypothesis that differences in eye movements during complex scene viewing of normal and colorblind individuals could be observed. The ultimate goal was to use the eye movements to covertly detect red-green color vision defects using complex scenes. Regions of interest (ROIs) were defined as salient areas derived from the method described in Chapter 3. The +ROIs and –ROIs were predicted areas of potential fixation targets for normal and colorblind individuals respectively. It was hypothesized that normal individuals would have a higher chance to fixate on the +ROIs than colorblind individuals. In contrast, it was also predicted that colorblind individuals would have a higher chance to fixate on the –ROIs than normal individuals. To test the eye movement hypotheses, the percentage of fixations on the ROIs, average fixation duration, the total number of fixations, and the average saccade length were measured.

Consistent with the prediction, normal individuals looked more often at the +ROIs and spent more dwelling time on the +ROIs than did the colorblind individuals. The percentage of fixations made on the +ROIs by the normal group was significantly higher than the chance estimates and higher than that made by the colorblind group. This result suggested that normal participants had higher tendencies to fixate at the red-green salient regions as compared to the colorblind individuals. Longer average fixation duration on the +ROIs of the normal participants means they spent more time than colorblind participants processing the same visual information in these areas. It can be explained that the +ROIs were more obvious to the normal participants than to the colorblind ones. The color feature of saliency model could identify color regions (e.g., +ROIs) from complex scenes that capture normal individual's attention better than colorblind individuals.

This result also supported several studies' argument that colorblind individuals have lower ability than normal individuals to identify some colored objects in complex surroundings [10, 13, 24, 97]. Figure 4.3 demonstrates examples of fixation patterns for normally sighted and a colorblind individuals. The example shows the variety of fixation patterns that are possible.

The colorblind participants gazed more often on the –ROIs than just by chance alone and more often than the normal participants. It is likely that the –ROIs attracted more attention from the colorblind individuals than from the normal individuals. However, their visual attention was not entirely dominated by the influence of the –ROIs. As they fixated on the –ROIs, it was found that their dwelling time on these regions was not longer than on the +ROIs. Interestingly, they also spent some time fixating on the +ROIs as indicated by their percentage of fixations on these areas which was higher than the chance factor. It appears

that, although the –ROIs were more obvious to the colorblind participants than to the normal participants, the –ROIs might not be the only factors that capture colorblind individuals' attention from complex scenes. Other factors such as intensity and contrast also guide attention.

(a) (b)

Figure 4.3. Example of fixation patterns (a) from a normal participant (b) from a colorblind participant. The +ROIs are represented with white circles.

4.5.1 Color, Intensity, and Orientation Channel Correlation

As discussed in last section, it was surprising to find that colorblind individuals fixated +ROIs at greater than chance levels. Colorblind individuals have less visual ability to see the red-green contrast regions as represented by the +ROIs when compared to the normally sighted individuals [13, 24]. Their observed fixations on the +ROIs could possibly be explained by two reasons.

First, it was suspected that some colorblind participants in the study might have mild red-green colorblindness. Those who have mild red-green deficiencies such as the anomalous trichromacy will perceive contrast due to the bright shades between red and green. In such cases, their visual search behavior may not differ greatly from normal individuals because of the influence of +ROIs on their visual attention. Second, color contrast of the +ROIs might be correlated with other feature contrast in the image. Colorblind participants who could not distinguish shades of red and green might be attracted by intensity or orientation contrast presented in the +ROIs. To further probe the correlation of image features, the saliency values from color, intensity, and orientation maps were investigated

Feature correlation was defined as the correlation between the saliency values of two features (e.g., color vs. intensity or color vs. orientation) for each image. This was measured by correlating saliency values of each pixel from one feature map with those in the corresponding pixel from the second feature map. The correlation *r* between two features of all images is calculated from

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2}\sqrt{\sum_i (y_i - \bar{y})^2}} \tag{4.1}$$

where x_i and y_i are the saliency values of corresponding pixels in the two feature maps, and $\bar{x}$ and $\bar{y}$ are the mean saliency values of pixels in the two feature maps.

It was found that correlation between color and intensity contrast was significant (r = 0.32, p = 0.038). On the other hand, correlation between color and orientation contrast was not significant (r = 0.20, p = 0.142).

This result implies that some colorblind participants might be attracted to +ROIs not by color contrast but by intensity contrast as compared to the adjacent areas of the image. Intensity contrast should contribute to fixation point selection made by colorblind observers. When the color is less salient due to defective cone response in colorblind individuals, color is no longer a main factor for determining attention [67]. Visual attention of some colorblind individuals could be captured involuntarily by other salient features, which could be, in this case, the intensity of the ROIs [1]. On the other hand, weak association between the +ROIs derived by color and orientation features implies that colorblind participants were not likely to fixate the +ROIs of complex scenes on the basis of image orientation.

4.5.2 Color as Visual Attractor

The aim of this section was to quantify observed fixations on the ROIs by shades of colors. By matching observed fixations on ROIs and then identifying colors, this would provide some quantitative evidence on which color the ROIs tend to be. Some studies suggested that red attracts more visual attention from normal people more than other colors [104, 105] whereas yellow is the most distinctive color to the majority of red-green colorblind individuals [3]. Accordingly, from 100 random selected photographs, it was predicted that the ROIs frequently viewed by normal individuals would be in shades of red

color. On the other hand, the ROIs frequently viewed by colorblind individuals would be in shades of yellow color.

To acquire colors of fixation location on the ROIs, first, observed fixation locations on the ROIs were overlaid on the current image to determine pixel coordinates. The matrix of RGB coordinates of corresponding pixels was then extracted. After the RGB data for all pixels of fixation locations on the ROIs was acquired, each RGB dataset was matched with the X11 color name dictionary (see [106]) to identify color name. Then the color name was classified into shades using color name mapping guideline (see [107]) to represent shades of color at each fixation location. For example, a fixation located on a +ROI contained the RGB values of 170;45,93. From X11 color lookup, this dataset was approximately close to maroon (178;46;96), which is under shade of reds according to color name mapping guideline. Therefore, this fixation location was classified as red. Figure 4.4 displays a sample of color name mapping with shades of green, red, and yellow colors.

With the use of RGB to color name mapping guideline, colors of all 1,129 fixation locations on +ROIs from 20 participants could be quantified into shades of reds, oranges, and greens whereas all 460 fixations on –ROIs were quantified into shades of yellows, browns, and blues. Figure 4.5 demonstrates the observed fixations on the ROIs classified by colors. Red shades were matched most often by the fixation on ROIs by normal participants. Yellow shades were matched most often by the fixation on ROIs by colorblind participants. Consistent with the prediction, it was found that in this study, the ROIs of normal individuals tended to be in red shades whereas the ROIs of colorblind individuals tended to be in yellow shades.

Color Name	RGB	RGB Hex	Shades of Green
Dark Green	47;79;47	2F4F2F	
DarkGreen	0;100;0	006400	
dark green copper	74;118;110	4A766E	
DarkOliveGreen1	202;255;112	CAFF70	
DarkOliveGreen2	188;238;104	BCEE68	
green	0;128;0	008000	
lime	0;255;0	00FF00	
green2	0;238;0	00EE00	

Color Name	RGB	RGB Hex	Shades of Red
DeepPink	255;20;147	FF1493	
IndianRed1	255;106;106	FF6A6A	
IndianRed2	238;99;99	EE6363	
firebrick	178;34;34	B22222	
firebrick1	255;48;48	FF3030	
firebrick2	238;44;44	EE2C2C	
firebrick3	205;38;38	CD2626	
firebrick4	139;26;26	8B1A1A	

Color Name	RGB	RGB Hex	Shades of Yellow
BlanchedAlmond	255;235;205	FFEBCD	
DarkGoldenrod1	255;185;15	FFB90F	
DarkGoldenrod2	238;173;14	EEAD0E	
DarkGoldenrod3	205;149;12	CD950C	
LightGoldenrod	238;221;130	EEDD82	
yellow	255;255;0	FFFF00	
yellow2	238;238;0	EEEE00	
yellow3	205;205;0	CDCD00	

Figure 4.4. Example of RGB to color name mapping ([107]). RGB values of fixation locations were classified into corresponding shades of colors.

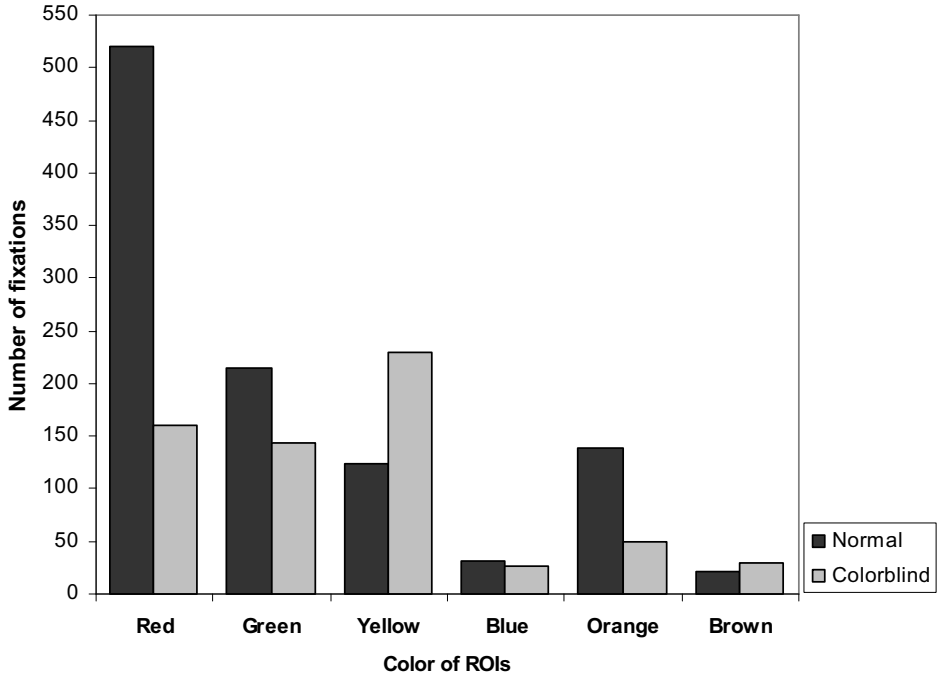


Figure 4.5. The number of fixations on ROIs of 100 images classified by shades of color.

4.5.3 Classification Performance

Because differences in the eye movements of the colorblind and normal individuals were observed when they viewed the selected complex images, the usefulness of this information as a performance metric for covert screening red-green color vision defects was considered. A similar approach as implemented in Chapter 2 was adopted to determine the efficacy of this metric. As an indication of the clustering success, the classification performance was developed by considering the percentage of fixations made on the +ROIs. The other eye movement measures (e.g., fixation duration, saccade length) were excluded from the classification performance analysis due to their wide range of observed values or relatively high variance in data. The wide range of observed values makes these eye movement variables more prone to misclassification.

Figure 4.6 displays the classification performance curve that plots percent correct as a function of the total number of fixations used in the analysis. To calculate the classification performance curve, it was assumed that there is an equal likelihood that normal and colorblind vision classes occurring out of male population (see Chapter 2; section 2.4.1.1). Every sequence of N fixations observed in the experiment was taken into account, where N ranged from 5 to 1125 fixations. First, f^*_{+ROI} was calculated for each participant. Then, it was assumed that the population of f^*_{+ROI} would be normally distributed $N(x|\mu,\sigma)$ with mean μ and standard deviation σ in order that population distributions could be estimated from the observed f^*_{+ROI} of the normal group (n) and the colorblind group (c). The unbiased classification threshold (x) was then determined as the intersection of two distributions where $N(x|\mu_n,\sigma_n) = N(x|\mu_c,\sigma_c)$. Classification performance was quantified as the percent of correct classifications using the unbiased classification threshold.

As can be seen in this figure 4.6, it requires 435 fixations to make a correct classification of color vision status in 95 percent of the cases. Given that the average number of fixations for each image was 15 fixations, 95 percent correct classification performance is achievable with the use of 29 images.

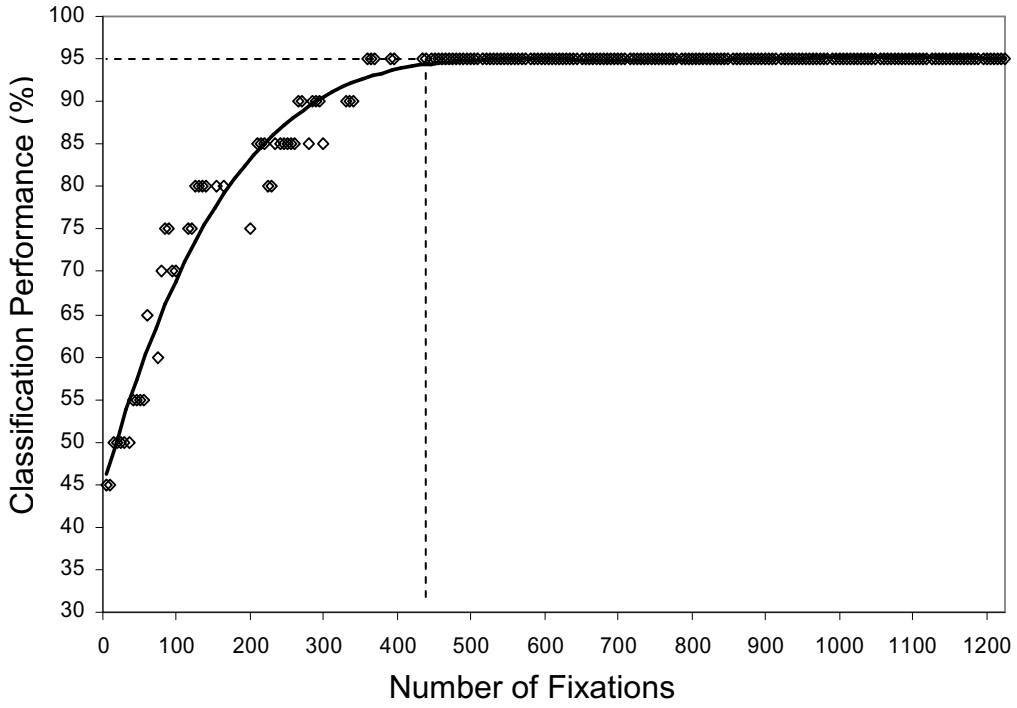


Figure 4.6. The classification performance curve indicates that approximately 435 fixations on selected scenes are required to obtain 95 percent correct color vision status classification accuracy.

Figure 4.7 shows the distribution of classification threshold (x) as a function of number of fixations taken from the experiment. As the number of fixations (*N*) increases, the classification threshold appears approximately constant, approaching the value of 3.51. On the basis of average f^*_{+ROI} and *N* for selected digital images, this result implies that if an individual's average f^*_{+ROI} is lower than 3.51, that individual will be classified as a colorblind. On the other hand, if such an average f^*_{+ROI} is higher than 3.51, that individual will be classified as a normal.

However, similar to Chapter 2's study, the fact that approximately 8.5% of the population is colorblind should be considered when applying this classification analysis to a test population. As the probability that normal and colorblind vision classes occurring out of

a population is not equal, the classification performance would result in a higher false positive rate and a lower false negative rate for the screening test.

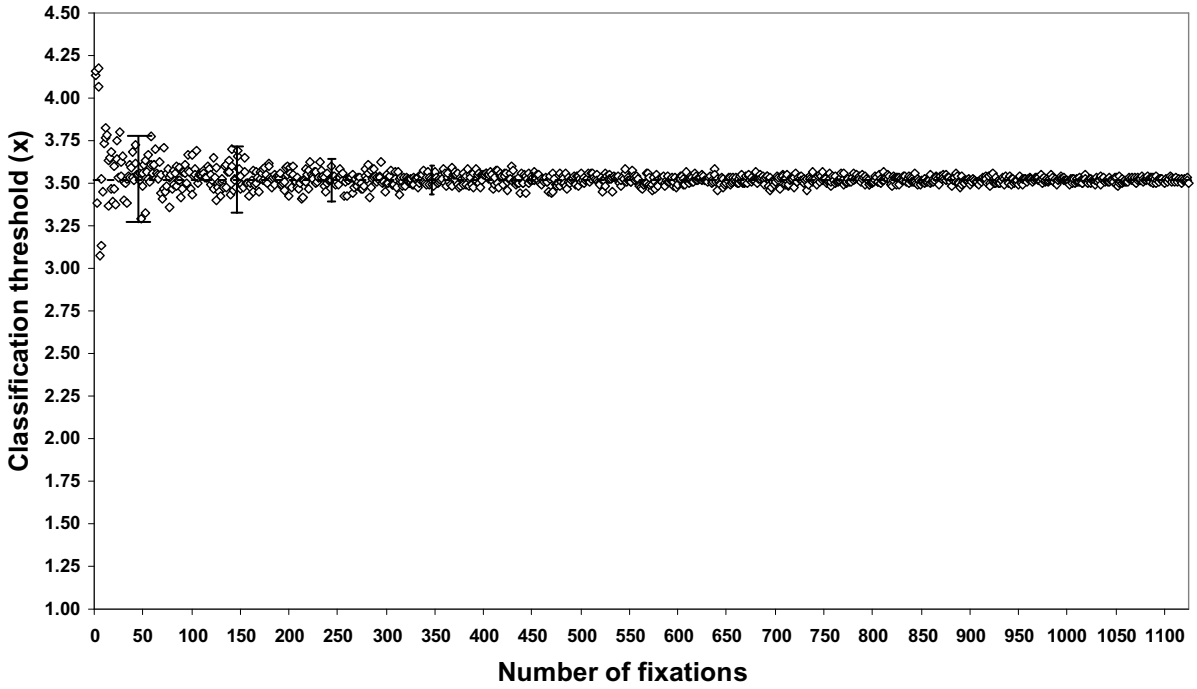


Figure 4.7. Relation between classification threshold (x) and number of fixations (*N*) on selected digital images. As number of fixations increases, the classification threshold appears approximately constant, approaching the value of 3.51. The error bar represents varying amplitude of the classification threshold at each interval between fixations.

4.5.4 Classification Performance Optimization

The +ROIs for each image were originally defined as the pixels that contained color saliency difference higher than 3 standard deviations ($\delta = 3$). However, $\delta = 3$ was selected based on an empirical rule that extreme values are generally higher than 3 standard deviations in the normal distribution. To achieve the rational balance of the classification, the amount of δ was adjusted to optimize the +ROIs, and the performance metric at each level of δ was considered.

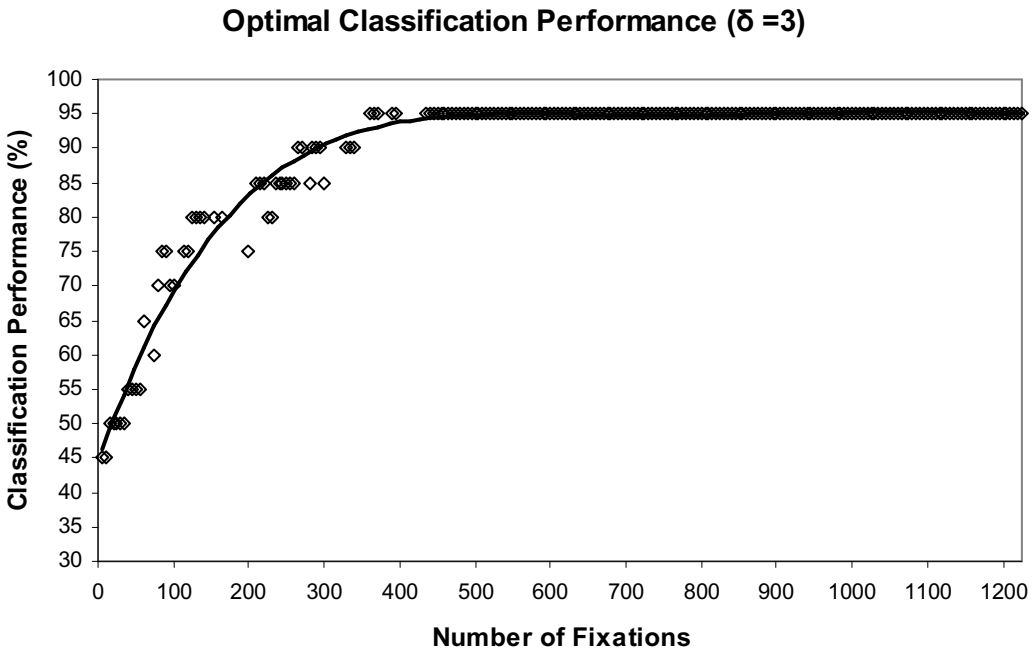
When the +ROIs are redefined by adjusting an amount of deviation, the number and area of the ROIs could be changed. If the δ is set too low, the size of +ROIs will be larger. Colorblind participants may fixate on these +ROIs although they are not really attracted by the +ROIs, per se. Colorblind participants would be classified as normal. This would increase the *"miss"* trials or failure to detect colorblind cases when in fact they were present (Type II error) .

On the other hand, if the δ is set too high, the size of +ROIs will be smaller. Normal participants' fixations may fall out of +ROIs, when in fact, they are influence by saliency in these regions. Normal participants would be classified as colorblind. Thus, this increases the risk of *"false alarm"* or overstating colorblind conditions (Type I error). In this case, it would take more fixations than usual to achieve any level of classification.

Figure 4.8 demonstrates the classification performance curve at different level of δ, using observed fixation from this study. When δ =1.5, a correct classification of color vision status could achieve only in 85 percent of the cases. A perfect classification was not possible with this data. In this case, the areas of +ROIs were considered large. Percentage of fixation on +ROIs of colorblind participants could be overestimated, thus inflating the miss cases.

When δ =2, the correct classification can achieve as high as 95 percent of the case with fewer number of fixations. When the δ =3, only 435 fixations were required to make a 95 percent correct classification of color vision status. However, as δ =3.5, approximately 1,050 fixations were required to make a 95 percent correct classification. In this case, the size of the +ROIs were small so that it took more fixations than usual to land on any +ROIs. Therefore, as δ is increased to a certain point, it requires more fixations than it should be to achieve the same level of correct classification.

(a)



(b)

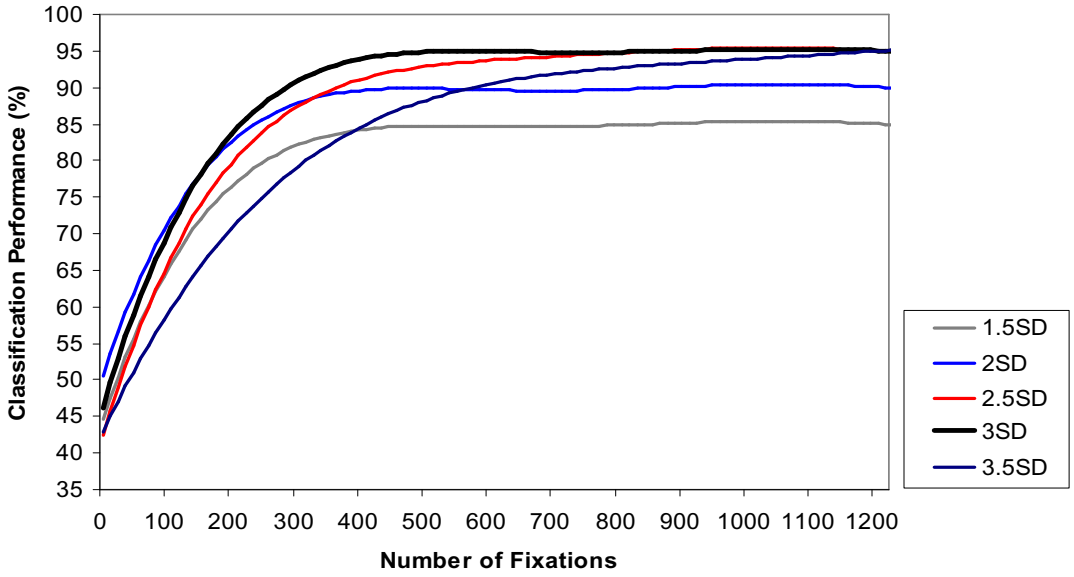


Figure 4.8. The classification performance curve optimization is shown with successive number of the fixations. (a) The optimal classifying performance with δ =3, requiring around 435 fixations to achieve 95%. (b) Five separate classification runs with different δ are shown from the same fixation data. Note that the required number of fixations to obtain a highest level of percent correct differs as the threshold for defining +ROIs changes from δ = 1.5 to 3.5. The optimal threshold value (δ =3) was used as representative curves for classifying color vision status.

CHAPTER 5. GENERAL CONCLUSION

5.1 Summary

This thesis was conducted to investigate the use of movement measurements to diagnose color vision deficiencies with three main objectives. The first objective was to extend the Ishihara testing method by using eye movements so that no active participation is required. The second objective was to develop a metric for screening color vision deficiencies using photographs to make the technique more covert. A method to select photographs and the region of interest useful for distinguishing eye movements of normal and colorblind individuals was developed. The third objective was to validate the proposed method by conducting an eye movement experiment in which normal and colorblind individuals were used as participants.

Known as color plates perceived differently by normal and colorblind observers, the Ishihara images were used in the first experiment. This study hypothesized that differences of eye movements between two groups should be observed. Ten normal and nine colorblind participants, first freely viewed the numerals represented in a series of 15 digital Ishihara images. In a second block of the experiment, participants were required to freely view and verbally report the numerals in each image. The results indicated that the eye movements of colorblind participants were characterized by significantly fewer fixations on the numerals, shorter fixation durations, and longer saccade lengths as compared to those of normal participants. As expected, correct identification rates for the numerals by the colorblind participants was low, but performance on individual images was positively related to the number of fixations made on the numerals during viewing.

To construct a classification metric for discriminating color vision status, all observed fixations on the numerals were used. It is noteworthy that 95 fixations were required to make a correct classification of color vision status in 95 percent of the cases. With the eye tracking implementation, only 3 Ishihara images are needed to present to participants to determine if they are classified into color normal or red-green colorblind group. To establish the validity of the test, the Cohen's Kappa coefficient of agreement (K) was used to account for the agreement of diagnostic results between the two tests. It was found that the agreement between the classification results of the computerized Ishihara test and the eye movement metric reached as high as 0.89 using as few as 5 images. In addition, the sensitivity and the specificity of the eye movement metric were comparable to those of the computerized Ishihara test. With approximately 95 fixations or 3 Ishihara images as classification criteria, the sensitivity and the specificity of the eye movement metric reached 0.88 and 1.0 respectively. It was shown that that the approach using eye movement metric on the Ishihara images in the first experiment fulfills the initial requirements of the objective to provide an alternative for a new color vision test.

One advantage of this test is that it is less susceptible to subterfuge due to the fact that it can be treated as a covert test. The individual being tested need not even be aware that they are being tested. This is an important advantage because participants might engage in subterfuge in order to pass the test for occupational reasons. While there exists only anecdotal evidence of subterfuge in color vision testing, with the introduction of color vision screening into the workplace for occupational safety reasons, professional organizations have become concerned with its possibility [4, 12].

While other color vision tests such as the arrangement test and the anomaloscope are difficult or impossible to cheat, such tests are more time consuming and require active participation on the part of the participant. This covert version of the Ishihara test is both shorter and easier than these other tests, and thus also applicable in mass testing scenarios. Mass testing could increase the numbers of individual screened, increase the number of deficiencies detected and could push the age of first diagnosis much earlier. Therefore, unlike other techniques, this covert color vision test can be uniformly applied to screen professional groups, elderly adults, children, or people with disabilities.

However, since the viewers can realize that they are being tested from viewing the Ishihara-like images, the second objective was established with the goal of using the selected complex scenes as stimuli with the objective to investigate the plausibility of covert color vision test. A method was developed to select photographs by applying the colorblind simulation model and the saliency model. This method takes random images as input and generates a subset of images that can distinguish visual attention of normal and color blind observers as output.

To summarize this method, after a random set of images with a predefined resolution is obtained, images with people, text, and numerals are manually filtered out. Images with no color are also removed. The Meyer and Greenberg's dichromatic algorithm then took the remaining images (e.g. normal images) as input and generate as output images with colorblindness simulation (e.g. colorblind images) to simulate perceived red-green color difference of colorblind observers. To acquire stimuli's color salience, both normal and the colorblind image pairs were processed with the saliency model to generate saliency maps [75]. The color feature saliency map of each image is used to emphasize color contrast,

marking the areas that normal and colorblind individuals tend to fixate. The subtraction of color saliency maps between the normal and colorblind image is calculated to acquire the color saliency difference. Skew of the color saliency difference is used as index to estimate the presence of clusters to attract visual attention differently from normally sighted and colorblind viewers. Images with the highest ranked skew are selected as stimuli. From the color saliency difference, the regions of interest for normal individuals (+ROIs) can be identified by the pixels that contained color saliency difference higher than δ standard deviations (+δSD) from the mean whereas the regions of interest for colorblind individuals (–ROI) are identified by the pixels that contained color saliency difference higher than δ standard deviations (–δSD) from the mean.

To validate this image selection method, the second experiment was hypothesized that the selected images could capture difference of eye movements between normal and colorblind participants. Normal participants were predicted to fixate more on the +ROIs while colorblind participants were predicted to fixate more on the –ROIs. Ten color normal and ten colorblind participants were instructed to freely view a series of 100 selected digital images. The results indicated that the eye movements of colorblind participants were characterized by significantly fewer fixations and shorter fixation durations on the +ROIs compared to those of normal participants. Colorblind individuals fixated on the –ROIs more often than normal participants.

To construct a classification performance metric, all observed fixations on the +ROIs were used, indicating that approximately 435 fixations or 29 selected scenes are required to obtain 95 percent correct of classification accuracy. In addition, the optimization analysis

suggests that defining the +ROI as $\delta = 3$ (e.g., 3 standard deviations from the color saliency difference) provided the rational balance for classifying color vision status.

In summary, the results from the studies show that observable differences in eye movements can be used as an efficient metric for detecting red-green defects with either the Ishihara images or the selected photographs. However, because individuals may become familiar with the Ishihara stimuli, photographs selected by the proposed method are suggested to be used as stimuli for a covert test rather than artificially constructed test images such as the Ishihara images.

5.2 Future Research

The purposes for future research on color vision test using eye movements are encouraging. As a new approach to screen individuals without their knowledge or active participation, it is expected that this covert color vision test could be applied to a wide range of situations. For example, this technique can be implemented as an enhanced forced-choice preferential looking paradigm on a computer screen. In this scenario, the position of correct response is randomly alternated by the computer, thereby reducing the possibility of observer's cheating or learning responses [12]. The eye movements are automatically recorded, thus minimizing human errors from the tester's judgment and eliminating requirements for user manual input.

With regards to the airplane accident previously discussed in Chapter 1, the scenario to apply this procedure for continually and timely testing pilot color vision without active participation could be researched. Using aforementioned forced-choice preferential looking paradigm implemented with an eye tracker, the instrument similar to an automated driver's

license vision screening system may be developed. In this case, it requires a pilot to look into the viewfinder at the correct color targets similar to the way the Department of Motor Vehicles (DMV) requires an applicant to read a series of numbers. The eye tracker then automatically analyzes the gaze movements and identifies the result. As a quick screening color vision test, this instrument might be used as a regular or random check point before a pilot enters the aircraft for higher aviation safety.

The covert test could also be used as a color vision screening at the point of a workstation with reference to visual perception inside the workplace. For example, a machinery station could automatically screen operators for colorblindness and halt operation upon detection of color vision abnormalities. For example, in a textile mill where color judgment in dyeing is critical, a computerized control system presents a random series of digital photographs of color fabrics. Simply by looking at these complex scenes displayed on a computer screen, the operator could be diagnosed with red-green color deficiency.

Another example of future research is the use of automatic testing in computer displays so that the graphical user interface can be adjusted if necessary for individuals with color deficits. In today's development of display systems and instruments, more color systems have been implemented since it provides more visual information as well as aesthetic attraction [108]. A few applications unfortunately have addressed the difficulties of colorblind users who may use these instruments on a regular basis or users who do not realize that they are colorblind. Eye movement measurement can be applied in this scenario as well. For example, a remote eye tracker can record eye movements while a user views a background image on a computer desktop or on the screensaver. This could be sufficient to covertly diagnose a color vision deficiency. The application could then generate a warning

with regards to color vision status or even provide a color adaptation mechanism (see [109]) on the color display to improve visual accessibility for colorblind users.

One drawback of this test is that the quality of eye tracking required necessitates either very expensive eye tracking hardware or a testing setup that restricts the movements of participant for the duration of the test. However, there is a significant amount of ongoing research aimed at lowering the intrusiveness and cost of eye tracking for the purposes of using eye movements in human computer interaction applications [110]. Furthermore, eye tracking technology is advancing rapidly. There is already technology to covertly track eye movements from over 10 meters away [111]. An eye tracker can be integrated into most computers and computer kiosks that are capable of recording user behaviors in the same way as digital cameras already are. Given recent advances, it is expected that within the next five years, these issues will be resolved, making widespread application of covert color vision testing a reality.

REFERENCES

1. Cole, B.L., *The handicap of abnormal colour vision.* Clin Exp Optom, 2004. **87**(4-5): p. 258-75.
2. Board, N.T.S., *Collision with trees on final approach Federal express flight 1478 Boeing Express Flight 1478 Tallahassee, Florida July 26,2002.* 2004: Washington, D.C.
3. O'Brien, K.A., et al., *Color and defective color vision as factors in the conspicuity of signs and signals.* Hum Factors, 2002. **44**(4): p. 665-75.
4. Spalding, J.A., *Colour vision deficiency in the medical profession.* Br J Gen Pract, 1999. **49**(443): p. 469-75.
5. Cole, B.L. and W.A. Macdonald, *Defective colour vision can impede information acquisition from redundantly colour-coded video displays.* Ophthalmic Physiol Opt, 1988. **8**(2): p. 198-210.
6. Engineers, E.P.D.o.t.I.o.E., *Color Vision Defects*. 2004, The Institution of Electrical Engineers: Hertfordshire.
7. Tagarelli, A., et al., *Colour blindness in everyday life and car driving.* Acta Ophthalmol Scand, 2004. **82**(4): p. 436-42.
8. Atchison, D.A., et al., *Traffic signal color recognition is a problem for both protan and deutan color-vision deficients.* Hum Factors, 2003. **45**(3): p. 495-503.
9. Kinney, J.A., H.M. Paulson, and A.N. Beare, *The ability of color defectives to judge signal lights at sea.* J Opt Soc Am, 1979. **69**(1): p. 106-10.
10. Birch, J., *Performance of red-green color deficient subjects on the Holmes-Wright lantern (Type A) in photopic viewing.* Aviat Space Environ Med, 1999. **70**(9): p. 897-901.
11. Hovis, J.K. and D. Oliphant, *Validity of the Holmes-Wright lantern as a color vision test for the rail industry.* Vision Res, 1998. **38**(21): p. 3487-91.
12. Group, S.R., *Minimum Color Vision Requirments for Professional Flight Crew - Part 1: The Use of Colour Signals and the Assessment of Colour Vision Requirements in Aviation*. 2006, Civil Aviation Authority: West Sussex. p. 2.
13. Cole, B.L. and K.Y. Lian, *Search for coloured objects in natural surroundings by people with abnormal colour vision.* Clin Exp Optom, 2006. **89**(3): p. 144-9.
14. LE, M. and J. RJ, *Electronic flight information system displays and colour defective observers.* Clinical and Experimental Optometry, 1991. **74**: p. 196-203.
15. Poole, C.J., et al., *Deficient colour vision and interpretation of histopathology slides: cross sectional study.* Bmj, 1997. **315**(7118): p. 1279-81.
16. Voke, J., *Colour vision defects--occupational significance and testing requirements.* J Soc Occup Med, 1978. **28**(2): p. 51-6.
17. Steward, J.M. and B.L. Cole, *What do color vision defectives say about everyday tasks?* Optom Vis Sci, 1989. **66**(5): p. 288-95.
18. Beard, B.L., W.A. Hisle, and A.J. Ahumada, *Occupational vision standards: A review*. 2002.
19. Committee on Vision, N.R.C., *Procedures for Testing Color Vision: Report of Working Group 41*. 1981, Washington, DC: The National Academy Press.

20. Kumagai, J.K., S. Williams, and D. Kline, *Vision Standards for Aircrew: Visual Acuity for Pilots*. 2005, Greenley and Associates Incorporated: Ottawa.
21. Spalding, J.A., *Medical students and congenital colour vision deficiency: unnoticed problems and the case for screening.* Occup Med (Lond), 1999. **49**(4): p. 247-52.
22. IET, *Colour vision defects: Colour blindness and Engineering*. 2006, Institution of Engineering and Technology: London.
23. L'Eclairage, C.I.d., *International recommendations for colour vision requirements for transport*. 2001, CIE: Vienna.
24. Dain, S.J., *Clinical colour vision tests.* Clin Exp Optom, 2004. **87**(4-5): p. 276-93.
25. Neitz, M. and J. Neitz, *Molecular genetics of color vision and color vision defects.* Arch Ophthalmol, 2000. **118**(5): p. 691-700.
26. Kinnear, P.R. and A. Sahraie, *New Farnsworth-Munsell 100 hue test norms of normal observers for each year of age 5-22 and for age decades 30-70.* Br J Ophthalmol, 2002. **86**(12): p. 1408-11.
27. Fernandez, R.H. and K.A. Turner, *Colour vision testing and the Electricity Supply Industry.* J Soc Occup Med, 1985. **35**(4): p. 137-9.
28. Adams, A.J., R. Balliet, and M. McAdams, *Color vision: blue deficiencies in children?* Invest Ophthalmol, 1975. **14**(8): p. 620-5.
29. Adams, A.J., et al., *Spectral sensitivity and color discrimination changes in glaucoma and glaucoma-suspect patients.* Invest Ophthalmol Vis Sci, 1982. **23**(4): p. 516-24.
30. (U.S.), C.t.R.t.S.E.o.t.P.N.R.C., *The polygraph and lie detection*. 2003, Washington, D.C.: National Academies Press. xvi, 398.
31. Pediatrics, A.A.o., *Use of photoscreening for children's vision screening.* Pediatrics, 2002. **109**(3): p. 524-5.
32. Teller, D.Y., *First glances: the vision of infants. the Friedenwald lecture.* Invest Ophthalmol Vis Sci, 1997. **38**(11): p. 2183-203.
33. Treisman, A.M. and G. Gelade, *A feature-integration theory of attention.* Cognit Psychol, 1980. **12**(1): p. 97-136.
34. Yantis, S., *Stimulus-driven attentional capture and attentional control settings.* J Exp Psychol Hum Percept Perform, 1993. **19**(3): p. 676-81.
35. Kowler, E., et al., *The role of attention in the programming of saccades.* Vision Res, 1995. **35**(13): p. 1897-916.
36. Parkhurst, D., K. Law, and E. Niebur, *Modeling the role of salience in the allocation of overt visual attention.* Vision Res, 2002. **42**(1): p. 107-23.
37. Parkhurst, D.J. and E. Niebur, *Texture contrast attracts overt visual attention in natural scenes.* Eur J Neurosci, 2004. **19**(3): p. 783-9.
38. Kaiser, P.K. and R.M. Boynton, *Human color vision*. 2nd ed. 1996, Washington, DC: Optical Society of America. xv, 652.
39. Schwartz, S.H., *Visual perception*. 2nd ed. 1999, Stamford, Conn.: Appleton & Lange. xiv, 433.
40. Fairchild, M.D., *Color appearance models*. 2nd ed. Wiley-IS&T series in imaging science and technology. 2005, Chichester, West Sussex, England ; Hoboken, NJ: J. Wiley. xxi, 385.
41. Livingstone, M., *Vision and art : the biology of seeing*. 2002, New York, N.Y.: Harry N. Abrams. 208.

42. Hubel, D.H., *Eye, brain, and vision*. Scientific American Library series ; no. 22. 1988, New York: Scientific American Library : Distributed by W.H. Freeman. viii, 240.
43. Marks, W.B., W.H. Dobelle, and E.F. Macnichol, Jr., *Visual Pigments of Single Primate Cones.* Science, 1964. **143**: p. 1181-3.
44. Brown, P.K. and G. Wald, *Visual Pigments in Single Rods and Cones of the Human Retina. Direct Measurements Reveal Mechanisms of Human Night and Color Vision.* Science, 1964. **144**: p. 45-52.
45. Bowmaker, J.K. and H.J. Dartnall, *Visual pigments of rods and cones in a human retina.* J Physiol, 1980. **298**: p. 501-11.
46. Dartnall, H.J., J.K. Bowmaker, and J.D. Mollon, *Human visual pigments: microspectrophotometric results from the eyes of seven persons.* Proc R Soc Lond B Biol Sci, 1983. **220**(1218): p. 115-30.
47. Mollon, J.D. and J.K. Bowmaker, *The spatial arrangement of cones in the primate fovea.* Nature, 1992. **360**(6405): p. 677-9.
48. Kremers, J., et al., *L/M cone ratios in human trichromats assessed by psychophysics, electroretinography, and retinal densitometry.* J Opt Soc Am A Opt Image Sci Vis, 2000. **17**(3): p. 517-26.
49. Walraven, P.L. and M.A. Bouman, *Fluctuation theory of colour discrimination of normal trichromats.* Vision Res, 1966. **6**(9): p. 567-86.
50. Nerger, J.L. and C.M. Cicerone, *The ratio of L cones to M cones in the human parafoveal retina.* Vision Res, 1992. **32**(5): p. 879-88.
51. Hurvich, L., *Colour Vision*. 1981, Sunderland, Mass.: Sinauer Associates.
52. Young, T., *On the theory of light and colours.* Philosophical Transactions of the Royal Society of London, 1802. **92**: p. 20-71.
53. Helmholtz, H., *On the theory of compound colours.* Philosophical Magazine, 1852. **4**: p. 519-542.
54. Wright, W.D., *A re-determination of the trichromatic coefficients of the spectral colours.* Transactions of the Optical Society, 1928. **30**: p. 141-164.
55. Guild, J., *The colorimetric properties of the spectrum.* Philosophical Transactions of the Royal Society of London, 1931. **A230**: p. 149-187.
56. Hering, E., *Outlines of a theory of the light sense*. 1964, Cambridge, Mass.,: Harvard University Press. xxvii, 317.
57. De Valois, R.L., I. Abramov, and G.H. Jacobs, *Analysis of response patterns of LGN cells.* J Opt Soc Am, 1966. **56**(7): p. 966-77.
58. Malacara, D., *Color vision and colorimetry : theory and applications*. 2002, Bellingham, WA: SPIE Press. vii, 164.
59. Westland, S. and C. Ripamonti, *Computational colour science using MATLAB*. 2004, Hoboken, NJ: J. Wiley. x, 207.
60. Gonzalez, R.C. and R.E. Woods, *Digital image processing*. 1992, Reading, Mass.: Addison-Wesley. xvi, 716.
61. Birch, J., *Efficiency of the Ishihara test for identifying red-green colour deficiency.* Ophthalmic Physiol Opt, 1997. **17**(5): p. 403-8.
62. Birch, J., *Diagnosis of defective colour vision*. 2nd ed. 2001, Boston: Butterworth-Heinemann. cm.

63. Dubuc, B., *The Brain from Top to Bottom: Vision*. 2002.
64. Egeth, H.E. and S. Yantis, *Visual attention: control, representation, and time course.* Annu Rev Psychol, 1997. **48**: p. 269-97.
65. Yarbus, A., *Eye movements and vision*. 1967, New York: Plenum Press.
66. Yantis, S. and J. Jonides, *Abrupt visual onsets and selective attention: evidence from visual search.* J Exp Psychol Hum Percept Perform, 1984. **10**(5): p. 601-21.
67. Theeuwes, J., *Top-down search strategies cannot override attentional capture.* Psychon Bull Rev, 2004. **11**(1): p. 65-70.
68. Desimone, R. and J. Duncan, *Neural mechanisms of selective visual attention.* Annu Rev Neurosci, 1995. **18**: p. 193-222.
69. Proulx, M.J. and H.E. Egeth, *Target-nontarget similarity modulates stimulus-driven control in visual search.* Psychon Bull Rev, 2006. **13**(3): p. 524-9.
70. Lamy, D., et al., *Effects of search mode and intertrial priming on singleton search.* Percept Psychophys, 2006. **68**(6): p. 919-32.
71. Folk, C.L., R.W. Remington, and J.C. Johnston, *Involuntary covert orienting is contingent on attentional control settings.* J Exp Psychol Hum Percept Perform, 1992. **18**(4): p. 1030-44.
72. Koch, C. and S. Ullman, *Shifts in selective visual attention: towards the underlying neural circuitry.* Hum Neurobiol, 1985. **4**(4): p. 219-27.
73. Itti, L. and C. Koch, *Computational modelling of visual attention.* Nat Rev Neurosci, 2001. **2**(3): p. 194-203.
74. Kelleher, J. and J. Genabith, *Visual Salience and Reference Resolution in Simulated 3-D Environments.* Artificial Intelligence Review, 2004. **21**(3): p. 253-267.
75. Itti L, K.C.K., Niebur E., *A Model of Saliency-Based Visual Attention for Rapid Scene Analysis.* IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 1998. **20**(11).
76. Maltz, M. and D. Shinar, *EYE MOVEMENTS OF YOUNGER AND OLDER DRIVERS.* Human Factors, 1999. **41**(1): p. 15-25.
77. Duchowski, A.T., *A breadth-first survey of eye-tracking applications.* Behav Res Methods Instrum Comput, 2002. **34**(4): p. 455-70.
78. Fukuda, R. and H. Bubb, *Eye tracking study on Web-use: Comparison between younger and elderly users in case of search task with electronic timetable service.* PsychNology Journal, 2003. **1**(3): p. 202-228.
79. Just, M.A. and P.A. Carpenter, *Eye fixations and cognitive processes.* Cognitive Psychology, 1976. **8**: p. 441-480.
80. Shepherd, M., J.M. Findlay, and R.J. Hockey, *The relationship between eye movements and spatial attention.* Q J Exp Psychol A, 1986. **38**(3): p. 475-91.
81. Rayner, K., *Eye movements in reading and information processing: 20 years of research.* Psychol Bull, 1998. **124**(3): p. 372-422.
82. Uttal, W.R. and E. Smith, *Recognition of alphabetic characters during voluntary eye movements.* Perception & Psychophysics, 1968. **3**: p. 257-264.
83. Li, D., *Low-cost eye-tracking for human computer interaction*, in *Human Computer Interaction*. 2006, Iowa State University: Ames.
84. Duchowski, A.T., *Eye tracking methodology : theory and practice*. 2003, New York: Springer. xvii, 251.

85. Cooke, L., *Eye Tracking: How It Works and How It Relates to Usability.* Technical Comminucation, 2005. **52**(4): p. 456-463.
86. Laboratories, A.S., *Eye tracking Instruction*. 1997, Applied Sciences Laboratories: Waltham, MA.
87. Ishihara, S., *The Series of Plates Designed as a Test for Colour-blindness, 3rd edition with 24 plates*. 1960: Kannehara Shubban, Tokyo.
88. Meyer, G. and P. Greenberg, *Color-defective vision and computer graphics displays.* Computer graphics & applications, 1988. **8**(5): p. 28-40.
89. Ing, E.B., J.A. Parker, and L.A. Emerton, *Computerized colour vision testing.* Can J Ophthalmol, 1994. **29**(3): p. 125-8.
90. Awad, Z., R.S. Natt, and D.D. Pothier, *Ishihara plates on your handheld computer.* Clin Otolaryngol, 2007. **32**(1): p. 58-9.
91. Hoffmann, A. and M. Menozzi, *Applying the Ishihara test to a PC based screening system.* Displays, 1998. **20**: p. 39-47.
92. Pardo, P.J., A.L. Perez, and M.I. Suero, *A new colour vision test in a PC-based screening system.* Displays, 2000. **21**: p. 203-206.
93. Shepherd, A.J., *Calibrating screens for continuous colour displays.* Spat Vis, 1997. **11**(1): p. 57-74.
94. Landis, J.R. and G.G. Kook, *The measurement of the observer agreement for categorical data.* Biometrics, 1997. **33**: p. 159-174.
95. Ramsey, F.L. and D.W. Schafer, *The statistical sleuth : a course in methods of data analysis*. 2002, Duxbury/Thomson Learning: Australia ; Pacific Grove, CA.
96. McPeek, R.M., V. Maljkovic, and K. Nakayama, *Saccades require focal attention and are facilitated by a short-term memory system.* Vision Res, 1999. **39**(8): p. 1555-66.
97. Spalding, J. and G. Arden, *Effects of Colour Blindness*. 2001, Surrey: Binfield Print and Design.
98. Abadi, R., *Book Review: Effects of Colour Blindness.* Ophthalmic and Physiological Optics, 2004. **24**(3): p. 252.
99. Okabe, M. and K. Ito, *How to make figures and presentations that are friendly to color blind people*. 2002, Japanese Drosophila Research Conference 2002.
100. rigden, c., *The eye of the beholder - designing for colour-blind users.* british telecomminucations engineering, 1999. **17**.
101. Rasche, K., R. Geist, and J. Westall, *Detail preserving reproduction of color images for monochromats and dichromats.* IEEE Comput Graph Appl, 2005. **25**(3): p. 22-30.
102. Anthony Nguyen, V.C., Sridha Sridharan, *Gaze tracking for region of interest coding in JPEG 2000.* Signal Processing: Image Communication, 2006. **21**: p. 359-377.
103. Alejandro Jaimes, J.P., Tim Grabowski, Jason Babcock, and Shih-Fu Chang. *Using Human Observers' Eye Movements in Automatic Image Classifiers*. in *SPIE conference on Human Vision and Electronic Imaging VI*. 2001. San Jose, CA.
104. Osberger W., R.A.M. *Automatic detection of regions of interest in complex video sequences*. in *Proceedings of Human Vision and Electronic Imaging*. 2001.
105. Drelie Gelasca, E.T., D. ; Ebrahimi, T. *Which Colors Best Catch Your Eyes: a Subjective Study of Color Saliency*. in *First International Workshop on Video*

Processing and Quality Metrics for Consumer Electronics, ISCAS. SPIE. 2005. Scottsdale, Arizona.
106. W3C, *CSS3 Color Module*. 2003.
107. Walsh, K., *RGB to Color Name Mapping*. 2007.
108. Firth, J., *Colour vision concerns in aviation, clinical concerns on colour coding in aviation.*, in *Operational Colour Vision in the Modern Aviation Environment, NATO RTO Technical Report 16.* 2001. p. 109-116.
109. Yang, S., Y. Ro, and J. Nam, *Improving Visual Accessibility for Color Vision Deficiency Based on MPEG-21.* ETRI, 2004. **26**(3): p. 195-201.
110. Li, D. and D. Parkhurst, *openEyes: An open-hardware open-source system for low-cost eye tracking.* Journal of Modern Optics, 2006. **53**(9): p. 1295-1311.

APPENDIX A. MATLAB CODE FOR DICHROMATIC SIMULATION

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Based on algorithm from "Gary W. Meyer and Donald G. Greenberg,
%Color-Defective Vision and Computer Graphics Displays,
%IEEE Computer Graphics and Applications, Volume 8 ,  Issue 5
%(September 1988), pp. 28-40"
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function dichromat(fname);
I = im2double(imread(strcat(fname,'.bmp')));

[m n r]=size(I);

XYZtoRGB=[3.2405 -1.5372 -0.4985; -0.9693 1.8760 0.0416; 0.0556 -0.2040
1.0573];
%XYZtolms=[0.3897 0.6890 -0.0787; -0.2298 1.1834 0.0464; 0.0000 0.0000
1.0000];
XYZtoSML=[0.0000 0.0000 0.5609; -0.4227 1.1723 0.0911; 0.1150 0.9364 -
0.0203];

RGBtoXYZ=inv(XYZtoRGB);

%RGBtoSML=RGBtoXYZ*XYZtoSML;

SMLtoXYZ=inv(XYZtoSML);
pcp=[0;0;1];    %protanopic confusion point in SML space
dcp=[0;1;0];    %deuteranopic confusion point in SML space
tcp=[1;0;0];

for i=1:m
    for j=1:n

        RGB= [I(i,j,1);I(i,j,2);I(i,j,3)];

        XYZ= RGBtoXYZ*RGB;
        %[Ixyz,Y]=XYZtoxyz(RGB);
        l=XYZ(2);%luminance
        [ua,va]=xyztouv(XYZ);

        %White Point D65 CIE1931
        w(1)=0.31271;
        w(2)=0.32902;
        w(3)=1-w(1)-w(2);
        [uw,vw]=xyztouv(w);


        %Transform for colorblind vision represented by deuteranopia
                DCPXYZ=SMLtoXYZ*dcp;    %deuteranopic confusion point in
XYZ space

                [up,vp]=xyztouv(DCPXYZ);%confusion point's u v value

                I477(1)=0.10278;
                I477(2)=0.10286;
                I477(3)=0.79436;
                [u477,v477]=xyztouv(I477);
```

```
                I578(1)=0.49915;
                I578(2)=0.49989;
                I578(3)=0.00096;
                [u578,v578]=xyztouv(I578);


                if (vp-va)*uw-(up-ua)*vw-(vp-va)*ua+(up-ua)*va<=0


                     u=((vw-va)*(u578-uw)-uw*(v578-vw)+ua*(vp-va)*(u578-
uw)/(up-ua))/((u578-uw)*(vp-va)/(up-ua)-(v578-vw));
                else

                     u=((v477-va)*(uw-u477)-u477*(vw-v477)+ua*(vp-va)*(uw-
u477)/(up-ua))/((uw-u477)*(vp-va)/(up-ua)-(vw-v477));

                end

                v=vp+(va-vp)*(u-up)/(ua-up);


        tempXYZ=[0;0;0];
        tempXYZ(2)=l;
        tempXYZ(1)=9.*u*l/(4.*v);
        tempXYZ(3)=((4./u-1.)*tempXYZ(1)-15.*l)/3.;

        tempRGB=XYZtoRGB*tempXYZ;

           while ((tempRGB(1)>1.)||(tempRGB(2)>1.)||(tempRGB(3)>1.))%hold
chromaticity and adjust luminance

                      tempXYZ=tempXYZ*0.99;

                   tempRGB=XYZtoRGB*tempXYZ;
           end

        for k=1:3
            I2(i,j,k)=clamp(tempRGB(k),0.,1.);
            %I2(i,j,k)=tempRGB(k);
        end

    end
end

%Create dichromatic image
imwrite(I2,strcat(fname,'_','d','.bmp'),'bmp');

function [u,v] = xyztouv(Iv);
        u=4.*Iv(1)/(Iv(1)+15.*Iv(2)+3.*Iv(3));
        v=9.*Iv(2)/(Iv(1)+15.*Iv(2)+3.*Iv(3));

function I=xyztoXYZ(Iv,l);
         I(1)=l/Iv(2)*Iv(1);
         I(2)=l;
         I(3)=l/Iv(2)*Iv(3);
```

```
function [I,l]=XYZtoxyz(Iv);
         l=Iv(2);
         I(1)=Iv(1)/sum(Iv);
         I(2)=Iv(2)/sum(Iv);
         I(3)=Iv(3)/sum(Iv);

function y=clamp(x,low,high);
        if x>high
            y=high;
        elseif x<low
            y=low;
        else
            y=x;
        end
```

APPENDIX B. 100 SELECTED DIGITAL IMAGES USED IN THE EXPERIMENT